



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 95244

TO: Manjunath N Rao
Location: cm1/10a11/10d01
Art Unit: 1652
Monday, June 09, 2003
Case Serial Number: 10/005499

From: Edward Hart
Location: Biotech-Chem Library
CM1-6B02
Phone: 305-9203

edward.hart@uspto.gov

Search Notes

Examiner Rao,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

THIS PAGE BLANK (USPTO)

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, .rnpm and .rnpn

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, .rapm and .rapn

The Pending database search results should not be left in the case because they contain data that is confidential.

THIS PAGE BLANK (USPTO)

95244

STIC-Biotech/ChemLib

Fr m: Rao, Manjunath N.
Sent: Friday, May 23, 2003 2:54 PM
T : STIC-Biotech/ChemLib
Subject: Sequence search request for 10/005,499

Fr m: Manjunath N. Rao
Art Unit 1652, Room 10A11
Mail Box in Room 10D 01
Phone: 306-5681

Date: 5-23-03

Please search the following as soon as possible for application with serial number **10/005499**

1. SEQ ID NO: 377 against all commercial nucleic acid databases including issued patents database and pending application database and provide a print of all results.

2. SEQ ID NO:378 against all commercial protein databases including issued patents database and pending application database and provide a print of all results.

If y u have any questions please call me at the above phone number.

Thanks

Manjunath N. Rao, Ph.D.
Bi technology Patent Examiner
Art Unit 1652, Room 10A11
Mail Box in 10D01
Crystal Mall 1, USPTO.

Edward Hart
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 6/2/03
Date Completed: 6/9/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:
NA Sequences: 1
AA Sequences: 1
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: 01/02
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library, CM1 – Circ. Desk



THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 16:28:18 ; Search time 5224 Seconds
(without alignments)
11231.098 Million cell updates/sec

Title: US-10-005-499-377

Perfect score: 2016

Sequence: 1 aggaaggttagggagcgga.....tggggctacagctgctgctg 2016

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_om:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vi:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_om:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_pat:*
- 24: em_ph:*
- 25: em_pl:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vi:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rtd:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1172.8	58.2	175186	2	AC026125	AC026125 Homo sapi
2	1172.8	58.2	194578	2	AC125611	AC125611 Homo sapi
3	1172.8	58.2	207607	2	AC010173	AC010173 Homo sapi
4	661	32.8	177172	2	AC096892	AC096892 Rattus no
5	661	32.8	179421	2	AC110690	AC110690 Rattus no
6	623	30.9	221470	2	AC027679	AC027679 Mus muscu
7	367.6	18.2	1526	9	BC008798	BC008798 Homo sapi
8	358.6	17.8	1297	9	AF095154	AF095154 Homo sapi
9	328.2	16.3	777	10	AF095155	AF095155 Mus muscu
10	305.8	15.2	2526	10	AB044560	AB044560 Mus muscu
11	296.8	14.7	970	10	BC024634	BC024634 Mus muscu
12	278	13.8	847	10	BC022724	BC022724 Mus muscu
13	271.6	13.5	1742	6	AX463516	AX463516 Sequence
14	259.2	12.9	981	6	AX456049	AX456049 Sequence
15	258.8	12.8	795	6	AX477323	AX477323 Sequence
16	237.8	11.8	750	6	AX477325	AX477325 Sequence
17	222	11.0	52065	9	AL353576	AL353576 Human DNA
18	220.4	10.9	195191	9	AC024339	AC024339 Homo sapi
19	207.8	10.3	198046	2	AL731670	AL731670 Mus muscu
20	205.4	10.2	10184	9	AF410771	AF410771 Homo sapi
21	205.4	10.2	204275	2	AC015936	AC015936 Homo sapi
22	203.4	10.1	11031	10	AB045983	AB045983 Mus muscu
23	157.2	7.8	81619	5	AL672065	AL672065 Zebrafish
24	152.2	7.5	185332	9	AC016673	AC016673 Homo sapi
25	151.4	7.5	94925	2	AC127906	AC127906 Rattus no
26	146.6	7.3	191804	2	AC019017	AC019017 Homo sapi
27	139.8	6.9	674	9	HS342389	HS342389 Homo sapi
28	133	6.6	190487	2	AC084310	AC084310 Mus muscu
29	131.4	6.5	94925	2	AC127906	AC127906 Rattus no
30	105.6	5.2	162599	9	AL360230	AL360230 Human DNA
31	100	5.0	1056	10	MUSBC1Q	M36293 Mouse compl
32	100	5.0	1059	6	AX305847	AX305847 Sequence
33	100	5.0	1059	10	MUSC1Q	M22531 Mouse compl
34	98.2	4.9	1136	6	AX401995	AX401995 Sequence
35	98.2	4.9	1136	10	RNC1QB	X71127 R.norvegicu
36	95.8	4.8	106037	2	AL158086	AL158086 Homo sapi
37	95.8	4.8	171553	2	AC013611	AC013611 Homo sapi
38	95.8	4.8	176558	2	AL611946	AL611946 Homo sapi
39	95.8	4.8	185848	2	AC025929	AC025929 Homo sapi
40	95.6	4.7	173239	2	AC113790	AC113790 Rattus no
41	94.2	4.7	977	9	HUMC1QB2	K03430 Human compl
42	93.8	4.7	182829	10	AL627214	AL627214 Mouse DNA
43	91	4.5	935	9	HSC1QBR	X03084 Human mrna
44	91	4.5	1076	9	BC008983	BC008983 Homo sapi
45	90.4	4.5	1357	6	AX068378	AX068378 Sequence

ALIGNMENTS

RESULT 1

AC026125

LOCUS

DEFINITION

Homo sapiens chromosome 12 clone RP11-234P5, WORKING DRAFT

SEQUENCE, 14 unordered pieces.

AC026125

AC026125.22 GI:21430995

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.

human.

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 175186)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Anaratunge,H.C., Are,J.R., Ayele,M., Banks,T.,

AC026125 175186 bp DNA linear HTG 18-JUN-2002
Homo sapiens chromosome 12 clone RP11-234P5, WORKING DRAFT
SEQUENCE, 14 unordered pieces.

Barbaria, J., Benton, J., Bimage, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P., Bunay, C., Burck, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.M., Earnheart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, J., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Hernandez, O., Hodgson, A., Hoques, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lozado, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, F., Martin, K., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokkenko, S., Oguh, M., Okwuonu, G., Oragunye, N., Oyedero, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Fu, L.L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojubokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N., Sisson, I., Sodergren, E., Sonalike, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabar, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villaion, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 175186)
Worley, K.C.
Direct Submission
Submitted (19-MAR-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 175186)
Worley, K.C.
Direct Submission
Submitted (18-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jun 17, 2002 this sequence version replaced gi:20336719.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Drafting Center Code: BCM
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: HAOU
Center clone name: RP11-234P5
----- Summary Statistics
Sequencing vector: M13;
Chemistry: Dye-primer Bodipy; % of reads
Chemistry: Dye-terminator Big Dye; 98% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 167945 bases at least Q40
Consensus quality: 170559 bases at least Q30
Consensus quality: 172253 bases at least Q20
Estimated insert size: 174551; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

```
OY 838 CCGCCAGCGCCCAAGGAGAGGTGGCGCGCGCGGAAAGCAGGCGCTCGGGGGCCCC 897
Db 49040 CCGCCAGCGCCCAAGGAGAGGTGGCGCGCGCGGAAAGCAGGCGCTCGGGGGCCCC 48981
OY 898 TGGACACACAGGTCCAGAGGGCCCCCAGGAGNACCCGCGCAGCCAGGCCCCCGGGGCC 957
Db 48980 TGGACACACAGGTCCAGAGGGCCCCCAGGAGNACCCGCGCAGCCAGGCCCCCGGGGCC 48921
OY 958 TCCCGGTTCAGGTTCGCGGGGGGTGGCGCGCGCTGCGGCTAGCTGCTCGCATTCCTTT 1017
Db 48920 TCCCGGTTCAGGTTCGCGGGGGGTGGCGCGCGCTGCGGCTAGCTGCTCGCATTCCTTT 48861
OY 1018 CTACGCGCGCGCTGCGCGCGCCCCACGAGGTTACGAGGTGCTGCGCTTCGACGACGTGT 1077
Db 48860 CTACGCGCGCGCTGCGCGCGCCCCACGAGGTTACGAGGTGCTGCGCTTCGACGACGTGT 48801
OY 1078 GACCAACGTGGCAACGCTACGAGCGCAGCCAGGCAAGTTACTTGTCCCAATGCCAGG 1137
Db 48800 GACCAACGTGGCAACGCTACGAGCGCAGCCAGGCAAGTTACTTGTCCCAATGCCAGG 48741
OY 1138 CGTCTACTTCTTCCGCTTACACGTGCTCATGCGCGCGCGCGCAGCCAGCATGTGGGC 1197
Db 48740 CGTCTACTTCTTCCGCTTACACGTGCTCATGCGCGCGCGCGCAGCCAGCATGTGGGC 48681
OY 1198 CGACCTCATGAGAACGACGACAGT 1221
Db 48680 CGACCTCATGAGAACGACGACAGT 48657
```

RESULT 4

AC096892/c

LOCUS

DEFINITION

Rattus norvegicus clone CH230-68F12, *** SEQUENCING IN PROGRESS

***, 74 unordered pieces.

AC096892

AC096892.5 GI:21728994

HTG: HTGS_PHASE1.

Norway rat.

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 177172)

REFERENCE

AUTHORS

Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Anaratunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbria,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrill,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgeson,A., Hoques,M., Holloway,C., Hollins,B., Homsl,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Loulseged,H., Lozadó,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Nowton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenwo,S., Ogih,M., Okunou,G., Oraqunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,

Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I., Sodergren,E., Sonaika,T., Sparks,A., Stanley,H., Stone,I., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstock,G. and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 177172)
Worley,K.C.

Direct Submission
Submitted (30-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 177172)
Worley,K.C.

Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:17972984.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GGWU
Center clone name: CH230-68F12
----- Summary Statistics
Sequencing vector: Plasmid:
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 114650 bases at least Q40
Consensus quality: 121322 bases at least Q30
Consensus quality: 126307 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 74 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1047: contig of 1047 bp in length
* 1048 1147: gap of unknown length
* 1148 2776: contig of 1629 bp in length
* 2777 2876: gap of unknown length
* 2877 4334: contig of 1458 bp in length
* 4335 4834: gap of unknown length
* 4835 5853: contig of 1419 bp in length
* 5854 5953: gap of unknown length
* 5954 7150: contig of 1197 bp in length
* 7151 7250: gap of unknown length
* 7251 8632: contig of 1382 bp in length
* 8633 10281: gap of unknown length
* 10281 10381: contig of 1549 bp in length
* 10382 12045: gap of unknown length
* 12046 12145: contig of 1664 bp in length
* 12146 13156: contig of 1011 bp in length
* 13157 13256: gap of unknown length
* 13257 14271: contig of 1014 bp in length
* 14271 14371: gap of unknown length
* 14371 15797: contig of 1427 bp in length
* 15798 15897: gap of unknown length


```
QY 829 GCCCCCTTCCCGCCAGGCGCCAAAGGAGAGGTGGCGCGCGCGGAAAGCAGGCGCTGCG 888
D 65078 GCCTCCCTTCCCTCCAGGTGCGCAAGAGAGAGTATAGGCGCGAGGAAAGCGCGCTGCG 85019
QY 889 GGGGGCCCTGGACCAACAGGTCCAGAGGGCCCCCAGGAGACCCGGGAGGCCAGGCC 948
D 65018 GGGACCCCGGAGGACCCCGGGTCCAGAGGGCCCCCTGGAGAGCCAGGAGGCCAGGTCC 64959
QY 949 CCGGGGCCCTCCCGGTCCAGGTCCGGGCGGGTGGCGCCCGCTGCGCGGTACGTGCTCG 1008
D 64958 TCCAGTCTCTCTGGCCCGAGCCCTGGAGGGCAGCGCCCTGCGCGGTATGTACCCCG 64899
QY 1009 CATTGCTTTTACGCGGGCGTGGCGGCGCCCGACAGAGGTTAGAGGTGCTGCGCTTCA 1068
D 64898 AATTGCTTTTATGCGGGCTAAGCGGCTCATGAGGGTTATGAGGTGTTGCGCTTGA 64839
QY 1069 CGAGTGGTGACCAAGCTGGCAACGCTTACGAGGAGCAGCGGCGCAAGTTTACTTGCC 1128
D 64838 CGAGTGGTGACCAAGCTGGCAACGCTTACGAGGAGCAGCGGCGCAAGTTTACTTGCC 64779
QY 1129 CATTGCCAGGCGTCTTCTTCTTACCAAGCTGCTCATGCGCGCGCGGCGGACCGACCG 1188
D 64778 CATTGCCAGGCGTCTTCTTCTTACCAAGCTGCTCATGCGCGCGGCGGAGCGGTACCG 64719
QY 1189 CATTGGGCGGACCTCATGAAGAACGAGCAGGTGCGG 1225
D 64718 CATTGGGCGGACCTCATGAAGAACGAGCAGGTGAGG 64682
```

RESULT 5

```
AC110690 LOCUS
DEFINITION Rattus norvegicus clone CH230-290M8, *** SEQUENCING IN PROGRESS
***, 29 unordered pieces.
AC110690 AC110690.4 GI:21740747
VERSION HTG; HTGS_PHASE1.
KEYWORDS Norway rat.
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 179421)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Anaratunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,B., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
Homsf., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Kovar,C.,
Li,J., Li,Z., Lichtarge,O., Liu,C., Liu,J., Liu,W., Lousegh,H.,
Lozadó,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,B., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokoko,S., Ogih,M., Okuwonu,G.,
Oragunye,N., Oyedoro,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Picken,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
```

```

Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shoostari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczyk,R., Woodson,S., Worley,K.,
Wu,C., Wu,Y., Wu,F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 179421)
Worley,K.C.
Direct Submission
Submitted (15-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 179421)
Worley,K.C.
Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:20514409.
-----
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: GRM
Center clone name: CH230-290M8
-----
Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 147907 bases at least Q40
Consensus quality: 151353 bases at least Q30
Consensus quality: 153945 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 29 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1305: contig of 1305 bp in length
* 1306: gap of unknown length
* 1406: contig of 1680 bp in length
* 3085: gap of unknown length
* 3185: gap of unknown length
* 4216: contig of 1031 bp in length
* 4316: gap of unknown length
* 4217: contig of 1579 bp in length
* 5895: gap of unknown length
* 5896: contig of 1011 bp in length
* 7007: gap of unknown length
* 7106: contig of 1700 bp in length
* 8806: gap of unknown length
* 10034: contig of 1128 bp in length
* 10134: gap of unknown length
* 11868: contig of 1734 bp in length
* 11968: gap of unknown length
* 13610: contig of 1642 bp in length
* 13710: gap of unknown length
* 13711: contig of 1375 bp in length
* 15085: gap of unknown length
* 15185: gap of unknown length
* 15186: contig of 1645 bp in length
* 16930: gap of unknown length
```


Qy	849	CCAAGGAGAGGTGGCGCGCGCGGGAAGACAGGCTTCGGGGGGCCCTCTGGACCAACCAG	908
Db	440	CCGAGGGAAGCGGGCGCACCGGAAGCCGGGCGCTCCGGGGCGCTCCCGGGAGACCCAG	499
Qy	909	GTCCAAGAGGGCC-----CCAGGAGAACCCGACGAGCCAGGCCGCC	950
Db	500	GTCTCTCCCGGCCCTGTGGGGCCCGGGGGAGAAGGTTGAGCCAGGACCGCGGGCCCTC	559
Qy	951	CGGGCCCTTCCCGTCCAGGTCCGGCGGGGTGGCGCCGCTGCCGGCTAC-----GTGC	1004
Db	560	CGGGGCTTCCCGGGGGGGGAGCGGGCGCCATCAGCACTGCCACCTACACACGGTGC	619
Qy	1005	CTCGCATGCTTTCTACGGGGGCTCGGGCGGCCCCACGAGGGTTAGGAGGTCTCGGCT	1064
Db	620	CGCGTGGGCTTCTACGCGGGGCTCAAGAAACCCACAGGGTTAGAGGTACTCAAGT	679
Qy	1065	TCGACGACGTGGTGACCAACGTGGGGAACGCTTACGAGGCAGCGACGGCAAGTTTACTT	1124
Db	680	TTGACGAGTGGTCACCAACCTAGGCAACAACATACGACGCGGCCAGGCTTTACGT	739
Qy	1125	GCCCATGCCAGCGTCTACTTCTTCCGCTTACCACGTGTCATGCGGGGGGAGCGGCA	1184
Db	740	GCACATTCCTCCGCACTTCTTTTCACTTACCATTGCTCTCATGCGCGGGCGACGGCA	799
Qy	1185	CCAGCATGTGGCGCGACCTCATGAAGAACGACAGTCCGGGGCAGCGCATTCGCTCAGG	1244
Db	800	CCAGTATGTGGCGAGACCTCTGCAAGNATGGCCAGGTGGGGCCAGTGCTATGCCCCAG	859
Qy	1245	ACGGGACACAGAACTACGAGTACGCAGCAACAGCGTCTATTCTGCACTCTGAGCGTGGGG	1304
Db	860	ACGCGACCAAGAACTACGACTACGCCAGCAACAGCGTGATCTCTGCACTCTGGACGCGGG	919
Qy	1305	ACGAGGTCTTCATCAAGCTGACGGCGGGAAAGTGACAGCGGCGCAACACCAACAGTACA	1364
Db	920	ACGAGGTCTTCATCAAGCTGATGGAGGCAAGCAACACGCGGCGACACAGCAACAATACA	979
Qy	1365	GCACCTTCTCCGGTTCAATCATCCCGACTGAGCGGGCCCGCCCGGTGCCCGGCT	1424
Db	980	GCACGTCTCTGGCTTCATCATCTACTCCGACTGAGTCCGCCACGCTCTCCTCCACCCAC	1039
Qy	1425	CGGCCCTTCTCTCCGTCCTCACCCACCTCTCTGCCGCCACCCGAGGGCGCACCCAC	1484
Db	1040	GTCCCTCACCGCGGGGTCCCTTCGGGGGGGAGACGATGACTGCGCCCTCGGCCAC	1099
Qy	1485	CTTTTGAGAGCCTGGC	1500
Db	1100	CCGCTGCTGCCGGC	1115

RESULT 8

AF095154
LOCUS AF095154 1297 bp mRNA linear PRI 15-OCT-1998
DEFINITION Homo sapiens C1q-related factor mRNA, complete cds.
ACCESSION AF095154
VERSION AF095154.1 GI:3747096

KEYWORDS
SOURCE
Homo sapiens.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
AUTHORS
1 (bases 1 to 1297)
Berube, N.G., Swanson, H.X., Bertram, M.J., Kittle, J.D., Didenko, V.,
Baskin, D.S., Smith, J.R. and Pereira-Smith, O.M.
TITLE
Cloning and characterization of CRF, a novel C1q-related factor,
expressed in areas of the brain involved in motor function
JOURNAL
Unpublished

REFERENCE 2 (bases 1 to 1297)

AUTHORS Berube, N.G., Swanson, H.X. and Perelra-Smith, O.M.

JOURNAL
TITLE
Direct Submission
 Submitted (28-SEP-1998) Huffington Center on Aging, Baylor College
 of Medicine, One Baylor Plaza, Houston, TX 77030, USA

FEATURES	
1	Location/Qualifiers
1	1287

source
T. :1297

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17q21"
14. .790
/note="CRF; contains a triple helical collagenous domain
and a C1q signature domain; highly expressed in the
brainstem"
/codon_start=1
/product="C1q-related factor"
/protein_id="AAG64186.1"
/db_xref="GI:3747097"
/translation="MLLVLVLPVLSGGPVGHEMLGTCRMVCDPVYPARGPGAGA
RTGGDALSQSGAPPSTLVQSPQSKPGRTGKPGPGDPGPGVPVGPQGEKEGP
KGKPGPLPAGGSGAISTATYITVPRVAYAGLKNPHEGTEVLKDFDDVVYINLGNND
AASGKFTCNIPGFTFYTHVLMRGDGTSMWDLCKNGQVRSALIAQDDAQNDQYDIAS
SVTLHLDGADVEPIKLDGGRAKGGNNKSYTFSGFIYSD"
218 a 484 c 406 g 189 t

```

[illegible]

869	QY	CCGCGGAAGACAGGCCTGGGGGGCCCCCTGGACACACAGGTCCCAAGAGAGGGCCCCCAGGA	928
241	Db	CCGGGGCCTCCAGGAGACGGGGGCCCTCCAGGTCTGTGGGCGCGCTGGGAGAGAAGGT	300
929	QY	GAACCCGACAGCCAGGCCCCCGGGCCCTCCCGTCCAGTCCCGGGCGGGGTGGCGCCC	988
301	Db	GAACACGAGAACCTGGTCTCTGGCTGCGGGTTCGGGAGGACGACGGCGGCATCAGC	360
989	QY	GCTGCGCGGTA-----CGTGCTCGCATTTCTACGCGGGCCTGCGGGGGCCCCAC	1042

361	ALGGGCAACCTATACACAGGTGCACAGCGTGGGCTTCTACGCGCGGCCCTCAAGGACGCCCTCAT	429
1043	GAGGGTTACGAGGTGTCGCCCTCGACGACGTGGTGACCAAGTGGCGACGCGCTACGAG	1102
QY		
421	GAGGGTTACGAGGTGCTCAAGTTTGACGACGTGGTGACCACTTAGGCAACAACACTACGAT	480
Dbb		
1103	GCAGCGCAGGGCGAAGTTTACTTTGCCCCATGCCAGGCGGTCTACTTCTTCGCTTACCAACGTG	1162
QY		

481	CGGGCAGCGGCAAGTTTACATGCAACATCTCGGCACCTACTTTTTCACCTACCAAGTC	540
Db		
1163	CTCATCGCGGGGGGACGGCACCGCATGTGGCGGACCTCATGAAGAACGGACAGGTC	1222
Qy		
541	CTCATCGCGGGGGGACGGCACCGCATGTGGCAGACCTCTGCAAGAACGGCCAGGTG	600
Db		
1223	CGGGCCAGGCCCATGTCTCAGACGGCGGACAGAACTACGACTACGCCAGCAACAGCGTC	1282
Qy		
601	CGGGCCAGTGCATTTGCCAGATGCAGACAGAACTACGACTATGCCAGCAACAGCGTG	660
Db		

1283	ATTTCGCACCTGGACGCTGGCGACAGAGTCTTCATCAAGCTGGACGCGGGAAGATGCAC	1342
QY		
Db		
661	ATCTTCGATCTGGACGCGGGGATCAGGTCTTCATCAAGCTGATGGAGCAAAAGCACAC	720
QY		
1343	GCGCGCAACACACAACAAAGTACAGCAGCTTCTCGGGTTCATCATCTACCCGACTGA	1399
QY		
721	GGGGGCAACAGCAACAACAATACAGCACTTTCTTGGTCTTCATCTACTCGCGATTGA	777
Db		

LOCUS	DEFINITION	2526 bp	mRNA	linear	ROD 03-OCT-2000
AB044560	Mus musculus mRNA for Gliacolin, complete cds.				
RESULT 10					
AB044560					
LOCUS					
721	GGGGGGACAGGACACAAATACACACATTTCTCGGCTTCCATCACTACCTCGGATGA				777

ACCESSION	AB044560	
VERSION	AB044560.1	GI:10566470
KEYWORDS		
SOURCE	Mus musculus	cdNA to mRNA.
ORGANISM	Mus musculus	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.	
	1 (sites)	
REFERENCE	Kolde, T., Aso, A., Yorihuzi, T. and Nagata, K.	
AUTHORS		
TITLE	Conformational requirements of collagenous peptides for recognition	

LINE	JOURNAL	REFERENCE
20428709	J. Biol. Chem. 275 (36), 27957-27963 (2000)	Watanabe, Y., Nagata, K. and Hosokawa, N.
2 (bases 1 to 2526)		

TITLE Direct Submission
JOURNAL Submitted (12-JUN-2000) Yoshihiko Watanabe, Institute for Frontier Medical Sciences, Kyoto University, Department of Molecular and Cellular Biology; Shougoin Kawara-cho 53, Kyoto City, Sakyou-Ku,

FEATURES
Location/Qualifiers
Kyoto Prefecture 606-8507, Japan
(E-mail: yotchi@frontier.kyoto-u.ac.jp, Tel: 81-75-751-3886,
Fax: 81-75-751-4646)

```
1. .2526
source
/organism="Mus musculus"
/db_xref="taxon:10090"
901. .1668
CDS
```

```
/note="Ctq-like gene expressed in Glial cells"  
/codon_start=1  
/product="Gliacolin"  
/protein_id="BAB15806.1".
```

BASE COUNT		629 a 653 c 664 g 580 t		ORIGIN	
Query Match 15.2%; Score 305.8; DB 10; Length 2526; Best Local Similarity 63.7%; Pred. No. 3.5e-34; Matches 541; Conservative 0; Mismatches 257; Indels 51; Gaps 3;					
QY	602	AGTGAGGAAGATTGTCGGCCCTGGCAGCGCTCGCGCTGAGCGCGCGCAAGAGGTTGGCG	661		
Db	820	AGACAAGAGATCGGCCCGCTGAGCCGCTCGTGGCAGCAGGAGCGCCTGATCGCC	879		
QY	662	GGCGCGCGCTGGAGTGGCCATGTGCTGCTGCTGCTGGTGGCCATCCCGTGCCTGGTG	721		
Db	880	GGCGGGCGCTGGGGTGGTGGTGGTCTTCTGCTGGTCATCTCATCCCGGTGGTG	939		
QY	722	CACAGTCCCGCGGGCCAGCGCACTACGAGATGCTGGGTGCTGCCGCACTGTGTGCGAC	781		
Db	940	AGTCTCGCGGCACGTCGCTCACTACGAGATGCTGGCACCTGCCCATGTTCTCGAC	999		
QY	782	CCGCATGGGCCCGTGGCCGTGGTCCGAGGCGCCCTGC	823	-----T	823
Db	1000	CCCTATGGGGGCACCAAGGCTCCACGACCGCGCCACTCCGACCGGGGCCCTTATGCAG	1059		
QY	824	TCCGTCGCCCTTCCCGCAGCGCGCAAGGAGAGTGGCGCGCGCGGAAAGCAGGC	883		
Db	1060	TCCTCGCCACTTTCATCCAGGTCCCAAGGCGAGCGGTAGGCGCGGGAAGCAGGC	1119		
QY	884	CTCGGGGGGCC-----CCTGGACCAACCAAGTCCAAAGAGGGCCCCCAGAGAACCC	934		
Db	1120	CCGCTGGGCCCGCGGAGAGCCCGACCGCGCGGCCCTGGGGCCCCCGGCGGAGAAG	1179		
QY	935	GGCAGCGCAGGCCCGCGGCCCTCCCGGTCCAGTTCGCGGGGGGGTGGCGCCGCTGCC	994		
Db	1180	GGAGACCGGGTCGCAAGGCTTTCGCCCGCGCCCTGGGGCGCCCGGCTGAACGGGGCC	1239		
QY	995	G-----GCTAGCTGCCCTCGCATGCTTCTTCTACGGGGGCGCTG	1030		
Db	1240	GGGGCTATCAGTGCACGCCACCTATAGCAGGGTGCCCAAGATCGCCTTTACGCTGGCCTC	1299		
QY	1031	CGCGCGCCACAGAGGGTTACGAGTGTCTGCGCTTCGACGACGTGTGTGACCAACAGTGGC	1090		
Db	1300	AAACGGCAGCAGGAAGCTACGAGTGTCTCAAGTTTCGACGACGCTGGTAACCAATCTCGGA	1359		
QY	1091	AACGCTACGAGCAGCCAGCGCAAGTTTACTTGGCCCATGCCAGGCGTCTACTTCTTC	1150		
Db	1360	AACCACTATGATCCCAACACGGGCAAGTTACCTGCTCCATACCGGGTATCTACTTCTTC	1419		
QY	1151	GCATTACACGTCTCATGGCGCGCGCAGCGCACCATGTGGCGCACCTCATGAAG	1210		
Db	1420	ACCTACCACTCTGATGCGCGAGGGAGCGCACCATGTGGCTGATCTCTGCAG	1479		
QY	1211	AACGGACAGTCCGGCGCAGCGCCATTGCTCAGGACCGGACCAAGCACTACGACTACGCC	1270		
Db	1480	AACAACAGGTGGTGTGTAGTGCATTTGCTCAAGATGCTGATCAGAAATACGACTATGCC	1539		
QY	1271	AGCAACAGGTCAATCTGCACCTGGAGGTGGCGAGAGGTCTTCATCAGCTGGACGCC	1330		
Db	1540	AGTAACAGTGGTCTCTACCTTGGAACACGAGAGTGAAGTCTATATCAAAATAGATGGT	1599		
QY	1331	GGGAAAGTCAGCGCGCAACACCAAGTACAGCACCTTCTCCGGCTTCATCATCTAC	1390		
Db	1600	GGGAAGCCACCGAGGAACACACAATACAGCACATCTCTCGAATTAATATTAT	1659		
QY	1391	CCCGACTGA	1399		
Db	1660	GCTGACTGA	1668		

RESULT	11
BC024634	970 bp mRNA linear ROD 07-AUG-2002
LOCUS	Mus musculus, clone MGC:25969 IMAGE:4240136, mRNA, complete cds.
DEFINITION	
ACCESSION	BC024634
VERSION	BC024634.1 GI:19353132
KEYWORDS	house mouse.
SOURCE	Mus musculus
ORGANISM	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
REFERENCE	1 (bases 1 to 970)
AUTHORS	Strausberg, R.
TITLE	Direct Submission
JOURNAL	Submitted (01-MAR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK	NIH-MGC project URL: http://mgc.nci.nih.gov
COMMENT	Contact: MGC help desk Email: cgapbs-f@mail.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center Center code: BCM-HGSC Web site: http://www.hgsc.bcm.tmc.edu/cdna/ Contact: amg@bcm.tmc.edu Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M., Yoon, V.S., Kovis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M., Richardson, S., Gibbs, R.A.
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAC Plate: 31 Row: f Column: 10 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.
source	Location/Qualifiers 1..970 /organism="Mus musculus" /db_xref="taxon:10090" /map=FVB/N /clone=MGC:25969 IMAGE:4240136" /tissue_type="Kidney, normal. 5 month old male mouse." /clone_lib="NCI_CGAP_Kid14" /lab_host="DH10B" /notes="Vector: pCMV-SPORT6" 18..785 /codon_start=1 /product="Unknown (protein for MGC:25969)" /protein_id="AAH24634.1" /db_xref="GI:19353133" /translation="MVLLLVLPVLVSAGTSAHYEMLTGRWCDPYGGTKKAPSTA ATFDRLMQSLPTFIQGPKEAGRPGKAGRPPGPPGVPGPEKGEGRCGLP GPGPALNGAIASIAISATSPKPIAFYAGLRKHGEYEVLRKDDVVYTNLGNHDPT GKTCSIPGYIFTFYHLVRMGSDGTSMADLKKNQVRSIAIQDADQNVDYASNSV LHLEPDGEVIKLDGGKAGHGNNKYSTFGFIYAD"
CDS	BASE COUNT 221 a 276 c 284 g 189 t ORIGIN
Query Match	14.7%; Score 296.8; DB 10; Length 970;
Best local similarity	65.2%; Pred. No. 7.9e-33;
Matches	511; Conservative 0; Mismatches 222; Indels 51; Gaps 3;
Qy	667 GGCGCTGGAGTGCCCATGCTGCTGCTGCTGTTGGCCCATCCGCTGCTGTTGCACAG 726
Dd	2 GGCGCTGGGGTGTGAATGCTGCTTCGCTGGTCACTCCTCATCCGGTGTGTTGACTC 61
Qy	727 CTCGCCGGGCCAGCGCACATACGAGATGCTGGGTGCTGCGCATGTTGTGCGACCGCA 786

Db 62 GCGCGGACGCTGCTCACTACGAGATGCTGGGACCTGCGCATGCTCTCGACCCCTA 121
Qy 787 TGGCGCCCGTGGCCCTGGTCCCGACGCGCGCCCTGC-----TTCGCT 828
Db 122 TGGGGGACCAAGGCTCCAGCAGCGCGCCCACTCGGACCGGGCCCTTATGCACTCCT 181
Qy 829 GCGCCCTTCCCGCAGCGCGCAAGGAGAGGTGGCGCGCGGGAAGAGAGCGCTGCG 888
Db 182 GCGGACCTTTCATCAGGCTCCAAAGCGAGCGCGGTAGCGCGGGAAGCGCGCGCG 241
Qy 889 GGGCGCC-----CCTGGACCAACAGGTCACAGAGGCGCGCCAGAGAACCGGAG 939
Db 242 TGGCGCCCGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGAG 301
Qy 940 GCGAGCGCCCGCGCGCGCTCCCGCTCCAGGTCGCGCGCGCGCGCGCGCGCGCGCG 995
Db 302 GCGGCTCGCAAGGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 361
Qy 996 -----GCTAGCTGCTTCGATGCTTTCACGCGCGCGCGCGCGCGCGCGCGCG 1035
Db 362 TATCAGTGCAGCCACCTATAGCAGGTCGCCCAAGATCGCCTTTTACGCTGCGCTCAAACG 421
Qy 1036 GCGGCGGAGGTTACGAGTGTCTGCGCTTCGAGCAGCTGCTGACCAACGTCGGCAACGC 1095
Db 422 GCAGCAGAGGCTACGAGTGTCTCAAGTTTCGACGAGCTGTGTATCAATCTCGGAAACCA 481
Qy 1096 CTACGAGCGCAGCGCGCGCAAGTTTACTTGCCCATGCGAGCGCTCTACTTCTTCGCTTA 1155
Db 482 CTATGATCCCAACGCGCGCAAGTTCACTGCTCCATACCGGCTATCTTCTTCACTCA 541
Qy 1156 CCAGCTGCTATGCGCGCGCGCGCGCACAGCATGTGGCGCGACCTCATGAAGACGG 1215
Db 542 CCAGCTGCTATGCGCGCGCGCGCGCACAGCATGTGGCGCTGCTCTCGAAACAA 601
Qy 1216 ACAGTCCGCGCAGCGCATGTCTCAGGCGCGGACAGCACTACGACTACGCGCAAG 1275
Db 602 CCAGTGGCTGTAGTCAAGTTGCTCAAGATGCTGATCAGAAATACGACTATGCGAGTAA 661
Qy 1276 CAGGCTCATCTGACCTGAGCTGGCGGAGGAGTCTTCATCACTGAGCGCGGGA 1335
Db 662 CAGTGGTCTTCCAGCTGGAAACAGGAGATGAAGTCTATATCAATTAGATGGGAA 721
Qy 1336 AGTCCAGCGCGCGCAACCAACAGTACAGCACCTTCTCGGCTTCATCTACCCCGA 1395
Db 722 AGCCCGGAGGAAACAAACAAATACAGCACATCTCTGATTTATTTATGCTGA 781
Qy 1396 CTGA 1399
Db 782 CTGA 785

RESULT 12
BC022724 847 bp mRNA linear ROD 07-AUG-2002
LOCUS Mus musculus, clone IMAGE:4503736, mRNA, partial cds.
DEFINITION BC022724
ACCESSION BC022724
VERSION BC022724.1 GI:18490546
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Strausberg, R.
Direct Submission
Submitted (01-FEB-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory

cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcdpaxil@stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LINL at: <http://image.llnl.gov>
Series: IRAC Plate: 44 Row: f Column: 10.
Location/Qualifiers
1. 847
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4503736"
/tissue_type="Eye, retina, mouse strain C57BL/6"
/clone_lib="NIH_MGC_94"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
<1..364
/codon_start=2
/product="Unknown (protein for IMAGE:4503736)"
/protein_id="AAH22724.1"
/db_xref="GI:18490547"
/translation="PHEGYEVLFKDDVVNLGNVYDAASGRFTCTNIPCTYFTYHVLV
RGDGTSMADLCKNGOVRAIAQADADQNDYASNSVILHLDAGDEVFINKDGGKAH
GGSNKKYSTFSGFIIYSD"
BASE COUNT 197 a 269 c 209 g 172 t
ORIGIN
Query Match 13.8%; Score 278; DB 10; Length 847;
Best Local Similarity 83.0%; Pred. No. 3.7e-30;
Matches 317; Conservative 0; Mismatches 63; Indels 0; Gaps 0;
Qy 1037 CCCACGAGGTTACGAGTGTGGCTTCGACGACGCTGGTACCAAGCTGGGCAACGCC 1096
Db 2 CCTCATGAGGTTACGAGTGTCTCAAGTTTGACGAGCTGGTCAACCACTAGGCAACAA 61
Qy 1097 TACGAGGAGCGGCGCAAGTTTACTTGGCCCATGCCAGGCTTACTTCTTCGCTTAC 1156
Db 62 TACGATGGCGCGGCGCAAGTTTACATGCACATCTCTGGCACCCTACTTTTACCTAC 121
Qy 1157 CACGTGCTCATGCGCGCGCGGCGGACGACGATGTGGCGGCTCATGAAGACGGA 1216
Db 122 CAGGTCTCTATGCGCGCGCGGCGGACGACGATGTGGCGGAGCTCTGCAAGAACGCG 181
Qy 1217 CAGGTCCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1276
Db 182 CAGGTCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 241
Qy 1277 AGCGTCAATTCGACCTGGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1336
Db 242 AGCGTCAATTCGACCTGGAGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 301
Qy 1337 GTGCACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1396
Db 302 GCACACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 361
Qy 1397 TGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1418
Db 362 TGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 383

RESULT 13
AX463516
LOCUS AX463516
DEFINITION Sequence 80 from Patent WO0248337.
ACCESSION AX463516
VERSION AX463516.1 GI:21886290
KEYWORDS human.
SOURCE

ORGANISM	Homo sapiens	unidentified
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	unclassified.
AUTHORS	1	1
TITLE	Duggan,B.M., Yao,M.G. and Griffin,J.A.	Su,E.W.
JOURNAL	Secreted proteins	Clq-related factor, homologous polypeptides and therapeutic uses
JOURNAL	Patent: WO 0248337-A 80 20-JUN-2002;	Patent: WO 0212475-A 1 14-FEB-2002;
JOURNAL	INCYTE GENOMICS INC (US)	BLI LILLY AND COMPANY (US)
FEATURES	Location/Qualifiers	Location/Qualifiers
source	1..1742	1..981
	/organism="Homo sapiens"	/organism="unidentified"
	/db_xref="taxon:9606"	/db_xref="taxon:32644"
BASE COUNT	365 a 543 c 555 g 279 t	55..732
ORIGIN		/note="PRIMATE-lp231"
		55..732
		/note="unnamed protein product"
		/codon_start=1
		/protein_id="CAD38249.1"
		/db_xref="GI:21715007"
		/translation="MQTSAPPLLLSSRYQGRSPKSGRSPKSGSETARPGWRPVG
		EKGDSARPALQLQITDGTAGVGVARGVGVGSEGVTSALGATGSGDCTGMAVVG
		LKSPHEVRLKFDVVTNIGNHYDPTTGKFSQVIRGIYFFVTHILMRGDCGTSMWAD
		LCKNGQVRASAIQAQADQNYDIASNVVHLDSGDEYVKLDGGRAGHGNKNNKSTYFS
		GFLLRPD"
		55..114
		sig_peptide
BASE COUNT	179 a 322 c 308 g 172 t	
ORIGIN		
		12.9%; Score 259.2; DB 6; Length 981;
		Query Match
		Best Local Similarity 70.8%; Pred. No. 1.7e-27;
		Matches 344; Conservative 0; Mismatches 142; Indels 0; Gaps 0;
QY	912 CAAGAGGGCCCCCAGGAGAACCCGCGCAGGCAGGCCCGCCCGCGCTCCCGGTCCAGGTC	971
DB	245 CCAGCGCGCTCGGGTGTGGCGCGCGCGCGGGTAGGTGGCGATCCGAGGGTGAAG	304
QY	972 CGGGGGGGTGGCGCGCGCTGCGCGCTACGTGCTGCGATTGCTTTCACGGGCGCTGC	1031
DB	305 TGACCAAGTGCCTGAGCGCCACCTTCAGCGCGGCCAACGATCGCTTCTATGTGGGTCTCA	364
QY	1032 GCGGCCCCCAGCAGGGTTACGAGGTGCTCGCTTCGACACGCTGGTGACCAACGTGGCA	1091
DB	365 AGAGCCCCCAGCAAGGCTATGAGTGTCTGAAGTTCGATGACGTGGTGACCAACCTCGGCA	424
QY	1092 AGCGCTACGAGCAGCAGCGCAGTATTTACTTGGCCCCATGCCAGCGCTCTACTTCTTCG	1151
DB	425 ATCACTATGACCCACCGCAAGTTCAGCTGCCAGTACGCGCATCTACTTCTTCA	484
QY	1152 CTTACACGTGCTCATGCGCGGGGAGCGGACACGACATGTGGGGCCGACCTCATGAAGA	1211
DB	485 CCTACCACATCTCATGCGCGGGGAGCGGACCGACGATGTGGGGCGGACCTCTGCAAGA	544
QY	1212 ACGGACAGCTCCGGCGCAGCCATGCTCAGAGCGCGGACCGACCACTACGACTAGGCCA	1271
DB	545 ACGGGAGGTCCGGGCGCAGGCCATTCACAGAGCGCCGACCAACTACGACTACGCCA	604
QY	1272 GCAACAGCTCTTCTGCACTGGAGCTGGGGGAGCAGGCTCTTCATCAAGCTGGAGCGG	1331
DB	605 GTACACAGCTGTGCTGCACTTGGATTTCAGGGGACGAACTGTATGTGAAGCTGGATGGG	664
QY	1332 GGAAGGTGACCGCGGCAACCAACAAAGTACAGCACCTTCTCCGGCTTCATCATCTACC	1391
DB	665 GGAAGGCTCAGCGGAGCAATAATAACAGTACAGCACGTTCTCTCGGCTTCTCTGTACC	724
QY	1392 CCGACT 1397	
DB	725 CCGATT 730	
RESULT 15		
LOCUS	AX477323	
DEFINITION	Sequence 1 from Patent WO0244373.	
ACCESSION	AX477323	
VERSION	AX456049.1	
KEYWORDS	unidentified.	
SOURCE		
LINEAR		
BP	981	
DNA		
PAT	06-JUL-2002	
AX456049		
LOCUS		
DEFINITION	Sequence 1 from Patent WO0212475.	
ACCESSION	AX456049	
VERSION	AX456049.1	
KEYWORDS	unidentified.	
SOURCE		
LINEAR		
BP	981	
DNA		
PAT	06-JUL-2002	
AX456049		
LOCUS		
DEFINITION	Sequence 1 from Patent WO0212475.	
ACCESSION	AX456049	
VERSION	AX456049.1	
KEYWORDS	unidentified.	
SOURCE		
LINEAR		
BP	981	
DNA		
PAT	06-JUL-2002	
AX456049		
LOCUS		
DEFINITION	Sequence 1 from Patent WO0212475.	
ACCESSION	AX456049	
VERSION	AX456049.1	
KEYWORDS	unidentified.	
SOURCE		
LINEAR		
BP	981	
DNA		
PAT	06-JUL-2002	
AX456049		
LOCUS		
DEFINITION	Sequence 1 from Patent WO0212475.	
ACCESSION	AX456049	
VERSION	AX456049.1	
KEYWORDS	unidentified.	
SOURCE		
LINEAR		
BP	981	
DNA		
PAT	06-JUL-2002	
AX456049		
LOCUS		
DEFINITION	Sequence 1 from Patent WO0212475.	
ACCESSION	AX456049	
VERSION	AX456049.1	
KEYWORDS	unidentified.	
SOURCE		
LINEAR		
BP	981	
DNA		
PAT	06-JUL-2002	
AX456049		
LOCUS		
DEFINITION	Sequence 1 from Patent WO0212475.	
ACCESSION	AX456049	
VERSION	AX456049.1	
KEYWORDS	unidentified.	
SOURCE		
LINEAR		
BP	981	
DNA		
PAT	06-JUL-2002	
AX456049		
LOCUS		
DEFINITION	Sequence 1 from Patent WO0212475.	
ACCESSION	AX456049	
VERSION	AX456049.1	
KEYWORDS	unidentified.	
SOURCE		
LINEAR		
BP</		

Search completed: June 6, 2003, 20:04:01
Job time : 5240 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 16:23:28 ; Search time 1334 Seconds

(without alignments)
3403.320 Million cell updates/sec

Title: US-10-005-499-377

Perfect score: 2016

Sequence: 1 aggaaggttaggagggcga.....tgggctacagctgctg 2016

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	DB	ID	Description	
1	717	35.6	717	24	ABK35597	Gene encoding novel	
2	620	30.8	804	24	ABK35598	Gene encoding novel	
3	305.8	15.2	2526	22	AAH48068	Murine HSP47 inter	
c	4	304.6	15.1	639	22	AAH21084	Parallel detection
5	303.2	15.0	768	24	ABA90355	Human polynucleoti	
c	6	286.4	14.2	584	24	ABQ45772	Oligonucleotide fo
7	286.4	14.2	584	24	ABQ45773	Oligonucleotide fo	
8	272.4	13.5	864	24	AAI57577	DNA encoding novel	
9	269	13.3	466	24	ABL78241	Human ovarian canc	

10	262.8	13.0	584	24	ABQ45770	Oligonucleotide fo	
c	11	262.8	13.0	584	24	ABQ45771	Oligonucleotide fo
12	259.2	12.9	981	24	ABK10666	Human CRF-like pro	
c	13	107.6	5.3	505	24	ABQ29322	Oligonucleotide fo
14	107.6	5.3	505	24	ABQ29323	Oligonucleotide fo	
c	15	101.4	5.0	505	24	ABQ29324	Oligonucleotide fo
16	101.4	5.0	505	24	ABQ29325	Oligonucleotide fo	
c	17	100.5	5.0	1059	24	ABQ199588	Mouse ischaemic co
18	98.2	4.9	1136	24	ABK63764	Rat sequence diffe	
c	19	95.2	4.7	15558	22	AAK39926	Genomic sequence #
20	95.2	4.7	15558	22	AAK90370	Human digestive sy	
c	21	93.2	4.6	600	22	ABA59702	Human foetal liver
22	93.2	4.6	600	22	AAK07971	Human brain expres	
c	23	93.2	4.6	600	22	AAK33841	Human bone marrow
24	93.2	4.6	600	22	AAI39567	Probe #8253 used t	
c	25	93.2	4.6	600	24	ABK08688	Human genome-deriv
26	91.8	4.6	987	22	AAK28263	Monkey MANGO 245 O	
c	27	91.8	4.6	1388	22	AAK28262	Monkey MANGO 245 D
28	91	4.5	997	22	AAI59763	Human polynucleoti	
c	29	91	4.5	1185	22	AAI57977	Human polynucleoti
30	90.4	4.5	1357	22	AAK28672	Human zacrpa codin	
c	31	87.8	4.4	1338	22	AAK44999	Human secreted pro
32	87.8	4.4	1338	22	AAK45000	Human secreted pro	
c	33	87.8	4.4	1338	22	AAK45001	Human secreted pro
34	86.2	4.3	1338	22	AAK44970	Human TANGO 253 co	
c	35	86.2	4.3	1338	22	AAK44998	Human secreted pro
36	85.6	4.2	114955	20	AAK53491	Human adenosine A1	
c	37	85.4	4.2	728	22	AAK44994	Human secreted pro
38	85.4	4.2	728	22	AAK44996	Human secreted pro	
c	39	85	4.2	1347	20	AAK24684	Human adipocyte-sp
40	85	4.2	1377	20	AAK87258	cDNA clone encodin	
c	41	85	4.2	1377	20	AAK80052	Human PRO344 nucle
42	85	4.2	1377	21	AAK58626	Human PRO344 prote	
c	43	85	4.2	1377	21	AAK01241	Human PRO344 prote
44	85	4.2	1377	21	AAK46907	cDNA encoding nove	
c	45	85	4.2	1377	21	AAA49560	Human PRO344 cDNA

ALIGNMENTS

RESULT 1

ABK35597
ID ABK35597 standard; DNA; 717 BP.

XX AC ABK35597;

XX DT 08-MAY-2002 (first entry)

XX DE Gene encoding novel human secreted or membrane-associated protein #16.

XX KW Human; secreted protein; membrane-associated protein; hypertension;
KW inflammatory disorder; neurological disorder; haematopoietic disorder;
KW skeletal developmental disorder; growth abnormality; autoimmune disorder;
KW neurodegenerative disorder; nervous system disorder; bacterial infection;
KW peripheral myelinopathy; viral infection; cancer; obesity; diabetes;
KW hypotension; sexual development disorder; blood disorder; gene; ds.

OS Homo sapiens.

XX PN WO200204600-A2.

XX PD 17-JAN-2002.

XX PF 12-JUL-2001; 2001WO-US21985.

XX PR 12-JUL-2000; 2000US-218033P.

XX PR 21-AUG-2000; 2000US-226517P.

XX PA (SMIK) SMITHKLINE BEECHAM CORP.

XX PA (SMIK) SMITHKLINE BEECHAM PLC.

XX PA (GLAX) GLAXO GROUP LTD.

XX

PT Determining the degree of cytosine methylation in genomic DNA, useful
for diagnosis and prognosis, comprises selective hybridization of
amplicons from chemically treated DNA

XX

PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX

CC This invention describes a novel method for determining the degree of
methylation of a particular cytosine in a motif 5'-CpG-3', present in a
genomic sample of DNA. The sample is treated chemically to convert
cytosine (C) but not methylated C, to uracil, then part of the genomic
DNA that contains the target C is amplified to form a labeled amplicon.
The amplicon is hybridised to two classes, each with at least one
member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
and the degree of hybridisation to both classes is determined from the
label on the amplicon. From the ratio of labels hybridised to the two
classes of oligomers, the degree of methylation is calculated. The method
is used: (i) for diagnosis and/or prognosis of side effects of
therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
of the central nervous, cardiovascular, gastrointestinal and respiratory
systems etc., particularly by detecting mutations or single nucleotide
polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
types and for investigating cell differentiation. The method allows the
methylation status of many C residues to be determined simultaneously.
ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
method for determining the degree of cytosine methylation described in
the disclosure of the invention.

XX

SQ Sequence 584 BP; 100 A; 45 C; 207 G; 232 T; 0 other;

Query Match 14.2%; Score 286.4; DB 24; Length 584;

Best Local Similarity 76.6%; Pred. No. 2.3e-43;

Matches 389; Conservative 0; Mismatches 116; Indels 3; Gaps 3;

QY

1215 GACAGTCCGGCCGACGCGCATTTGCTCAGGACGCGGACGAGTACGACTACGCGACA 1274

DB

507 GATAAATCGAACCAACGCGCATTTACTCAAAACGCGAACCAGTACGACTACGCGACA 448

QY

1275 ACAGGCGTATCTGACCTTGGAGTGGGGGACGAGTCTTTCATCAAGCTGGACGCGGGA 1334

DB

447 ACAAGTCTATCTACACCTTAAACGTAAGCGGAAATCTTCATCAAACTTAAACGAGAA 388

QY

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

DB

387 AAATACACGACGACACCAACAATAACACACCTTCTCCGACTT-ATCATCTACCCG 329

QY

1395 ACTGAGCCGCGCCGCGCGCTGCGCCGCTTCTCTCCGCTTCTCTCCGACTT-ATCATCTACCCG 1452

DB

328 ACTAAACGACCGCCGCGCTGACCCGCTGCGCCGCTTCTCTCCGACTTCTCACCACCT 269

QY

1453 TCGTCCGCGCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1512

DB

268 TCTTAACCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 209

QY

1513 CTTTCGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1572

DB

208 CTTTCGCTTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 149

QY

1573 CCGGGAAGAGGAGGAGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1632

DB

148 CCGAANAANAANAANAACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 89

QY

1633 CCCAAGTGGCGCTGGACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1692

DB

88 CCCAATAGCGCTTAAACCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1

QY

1693 TCGGGAGGAGGTTTATAGCGGACTCA 1720

DB

28 TAAANAACAATTTTAAATACGAAGTCA 1

RESULT 7

ABQ45773

ID

XX

AC

XX

DT

XX

DE

XX

KW

KW

KW

KW

OS

XX

PN

XX

PD

XX

PF

XX

PR

XX

PR

XX

PA

XX

PI

XX

DR

XX

XX

PT

PT

XX

PS

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

XX

SQ

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

ABQ45773 standard; DNA; 584 BP.

ABQ45773;

12-JUL-2002 (first entry)

Oligonucleotide for detecting cytosine methylation SEQ ID NO 32364.

Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;

drug; side effect; cancer; central nervous system; cardiovascular;

gastrointestinal; respiratory system; single nucleotide polymorphism;

SNP; cell differentiation; ds.

Homo sapiens.

WO200218632-A2.

07-MAR-2002.

01-SEP-2001; 2001WO-EP10074.

01-SEP-2000; 2000DE-1043826.

05-SEP-2000; 2000DE-1044543.

(EPIG-) EPIGENOMICS AG.

Olek A, Piepenbrock C, Berlin K, Guetig D;

WPI; 2002-371829/40.

Determining the degree of cytosine methylation in genomic DNA, useful

for diagnosis and prognosis, comprises selective hybridization of

amplicons from chemically treated DNA

Claim 12; 56pp + Sequence Listing; 56pp; German.

This invention describes a novel method for determining the degree of

methylation of a particular cytosine in a motif 5'-CpG-3', present in a

genomic sample of DNA. The sample is treated chemically to convert

cytosine (C) but not methylated C, to uracil, then part of the genomic

DNA that contains the target C is amplified to form a labeled amplicon.

The amplicon is hybridised to two classes, each with at least one

member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers

and the degree of hybridisation to both classes is determined from the

label on the amplicon. From the ratio of labels hybridised to the two

classes of oligomers, the degree of methylation is calculated. The method

is used: (i) for diagnosis and/or prognosis of side effects of

therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders

of the central nervous, cardiovascular, gastrointestinal and respiratory

systems etc., particularly by detecting mutations or single nucleotide

polymorphisms (SNP's); and (ii) for differentiation of cell or tissue

types and for investigating cell differentiation. The method allows the

methylation status of many C residues to be determined simultaneously.

ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the

method for determining the degree of cytosine methylation described in

the disclosure of the invention.

Sequence 584 BP; 232 A; 207 C; 45 G; 100 T; 0 other;

Query Match 14.2%; Score 286.4; DB 24; Length 584;

Best Local Similarity 76.6%; Pred. No. 2.3e-43;

Matches 389; Conservative 0; Mismatches 116; Indels 3; Gaps 3;

1215 GACAGTCCGGCCGACGCGCATTTGCTCAGGACGCGGACGAGTACGACTACGCGACA 1274

78 GATAAATCGAACCAACGCGCATTTACTCAAAACGCGAACCAGTACGACTACGCGACA 137

1275 ACAGGCGTATCTGACCTTGGAGTGGGGGACGAGTCTTTCATCAAGCTGGACGCGGGA 1334

138 ACAAGTCTATCTACACCTTAAACGTAAGCGGAAATCTTCATCAAACTTAAACGAGAA 197

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

1335 AAGTCAGCGGGGACACCAACAGTACAGCAGCCTCTCCGCTTCATCATCTACCCG 1394

DE Human ovarian cancer related cDNA clone SEQ ID NO:1219.
KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.
OS Homo sapiens.
XX WO200192581-A2.
PN 06-DEC-2001.
XX 29-MAY-2001; 2001WO-US17756.
XX 26-MAY-2000; 2000US-207484P.
PR (CORI-) CORIXA CORP.
XX Algate PA, Harlocker SL, Jones R;
PI WPI; 2002-122075/16.
XX Composition for therapy and diagnosis of ovarian cancer comprising
PT polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
PT polypeptide, antibody specific to polypeptide or T cell expressing
PT polypeptide
XX
XX Claim 1; SEQ ID 1219; 489pp; English.
PS The present invention describes a composition (I) comprising: carriers
CC and immunostimulants; and a polypeptide (II) of a ovarian tumour
CC polypeptide encoded by a polynucleotide (III) having a cDNA sequence
CC (S1) from the 10912 nucleotide sequences as given in ABL77023 to
CC ABL87934, (III) encoding (II) having a sequence (S2), a T cell
CC population of (II), or antigen presenting cells that express (II).
CC (I) has cytostatic activity. An oligonucleotide (IV) that hybridises to
CC (S1) can be used for detecting ovarian cancer in a patient's biological
CC sample preferably serum or ovarian tissue. The method comprises
CC contacting a biological sample from a patient with (IV), detecting the
CC amount of polynucleotide hybridising to (IV) and comparing the amount to
CC a predetermined cutoff value and thereby detecting ovarian cancer in the
CC patient, where the amount of polynucleotide hybridising to (IV) is
CC detected preferably by polymerase chain reaction (PCR). (I) comprising
CC (III) and/or (II) is useful for stimulating and/or expanding T cells
CC specific for an ovarian tumour protein comprising contacting T cells
CC with (III) or (II). (III) is useful in design and preparation of
CC ribozyme molecules for inhibiting expression of the tumour polypeptides
CC and proteins in tumour cells; and to isolate a full length gene from a
CC suitable library e.g., a tumour cDNA library using well known
CC techniques.
XX
XX Sequence 466 BP; 110 A; 117 C; 126 G; 109 T; 4 other;
SQ
Query Match 13.3%; Score 269; DB 24; Length 466;
Best Local Similarity 96.8%; Pred. No. 3.4e-40;
Matches 338; Conservative 0; Mismatches 5; Indels 6; Gaps 6;
QY 1606 GAGCGACTCCGGAGGATACCCGACCCAGTCGCGCTGAGCCATAGGGCAGAG 1665
DB 1 GAGCGACTCCGGA-GGATACCCGACCCAGTG-GCGCTGAGCCCATAGGGCAGAG 57
QY 1666 GTCTGGCTTCTCTTTGTACAGAGATGGGAGCAGATTTTAAATCGGGACTCAGAGGC 1725
DB 58 GTCTGGCTTCTCTTTGTACAGAGATGGGAGCAGATTTTAAATCGGGACTCAGAGGC 117
QY 1726 CCAGAAAGCCGAGGAGCCCGCCAGCTTCGAGGGAATTAACAGAAACAGGAGGAGC 1785
DB 118 CCAGAAAGCCGAGGAGCCCGCCAGCTTCGAGGGAATTAACAGAAACAGGAGGAGC 177
QY 1786 CCATTAGCAAGAGAAGACATTAACAGAGGTAGTCAGGTCTCCGTCACAACCTTCT 1845
DB 178 CCATTAGCAAGAGAAGACATTAACAGAGGTAGTCAGGTCTCCGTCACAACCTTCT 237
QY 1846 CTCGCCACCTCTCGTCCCTCGTCTCCACTTCAGGCTTCAGGCTCCAGCCTTGGCAGGC 1905
|||||

DB 238 CTCGCCACCTCTCGTCCCTCGTCTCCACTTTTCAGGCTCAGGCTCCAG-CTTGGCAGCC 296
QY 1906 TTCCTGTGAACCT-GGAGGAACCACT-GAATTTCTTCTGCGCATTTAAAA 1952
|||||
DB 297 TTCCTGTGAACCTGGGAGGAACCACTGGAATTTCTTCTGCGCATTTAA 345
RESULT 10
ABQ45770
ID ABQ45770 standard; DNA; 584 BP.
XX AC ABQ45770;
XX 12-JUL-2002 (first entry)
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 32361.
DE Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
XX Homo sapiens.
OS
XX WO200218632-A2.
PN 07-MAR-2002.
XX 01-SEP-2001; 2001WO-EP10074.
XX 01-SEP-2000; 2000DE-1043826.
PR 05-SEP-2000; 2000DE-1044543.
XX (EPIC-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K, Guetig D;
PI WPI; 2002-371829/40.
DR Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA
XX Claim 12; 56pp + Sequence Listing; 56pp; German.
XX This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations of single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
XX Sequence 584 BP; 110 A; 45 C; 167 G; 262 T; 0 other;
SQ
Query Match 13.0%; Score 262.8; DB 24; Length 584;
Best Local Similarity 72.4%; Pred. No. 4.6e-39;
Matches 368; Conservative 0; Mismatches 137; Indels 3; Gaps 2;
QY 1215 GACAGTCCGGCCAGCGCCATTTGTCAGGACCGGACCGAGACTAGGACTAGCCAGCA 1274
|||||

Db 78 GATAGTTCGGGTTACGCTTATTTAGGACGCGGATTTAGAAATACGATTACGTTAGTA 137
QY 1275 ACAGCGTTCATCTGACCTGGACGTGGCGACGAGTCTTCATCAAGCTGGACGCGGA 1334
Db 138 ATAGCGTATTTGTTATTTGGACGTGGCGACGAGTCTTTATTAAGTTGGACGCGGA 197
QY 1335 AAGTCAGCGGCGCAACCAACAGTACAGACCTTCCGGCTTCATCATCTACCCG 1394
Db 198 AAGTCAGCGGCGTAATTAATAAGTATAGTATTTTTCGG-TTTATTTATTTTCG 256
QY 1395 ACTGAGCGGCG 1452
Db 257 ATGAGTCGGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCG 316
QY 1453 TCTCTCCG 1512
Db 317 TTTTGGTCGTTTATTCGAGCGGTTATTTATTTTGGAGATTGGCGGTGGAT 376
QY 1513 CTTCCGTTCCGGAGCGCGCTAAATGGCGAATCTTGTGCTCAAGGATTAAGTG 1572
Db 377 TTTTTCGTTTCGTTGGCGGTTAAATGGCGAATTTTGTGTTAAAGGTTAAGTGG 436
QY 1573 CGGGAAGGAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1632
Db 437 TCGGGAAGGAGGAGGAGTTCGGTTAGAGGAGTAGAGCGATTTCGGAGGAGTATTCGTA 496
QY 1633 CCCAAGTCGGCTCGACCCCATAGGGCAGAGGTCGTGGCTTCTCTTTTGTACAGAGA 1692
Db 497 TTTAAGTCGGCTGGATTTATAGGGGTAGAGTCGTGGTTTTCGTTTTCGTTATAGAGA 556
QY 1693 TGGGAGCAGTTTAAATAGCGGACTCA 1720
Db 557 TGGGAGTAGTTTAAATAGCGGATTA 584

RESULT 11
ID ABQ45771/c
AC ABQ45771; standard; DNA; 584 BP.
DT
TX
DE 12-JUL-2002 (first entry)
DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 32362.
KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
KW drug; side effect; cancer; central nervous system; cardiovascular;
KW gastrointestinal; respiratory system; single nucleotide polymorphism;
KW SNP; cell differentiation; ds.
OS
OS Homo sapiens.
PN WO200218632-A2.
PD 07-MAR-2002.
PF 01-SEP-2001; 2001WO-EP10074.
PR 01-SEP-2000; 2000DE-1043826.
PR 05-SEP-2000; 2000DE-1044543.
XX
XX (EPIG-) EPIGENOMICS AG.
XX Olek A, Piepenbrock C, Berlin K, Guetig D;
XX WPI; 2002-371829/40.
XX
XX Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA -
XX
PS Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
SQ Sequence 584 BP; 262 A; 167 C; 45 G; 110 T; 0 other;
Query Match 13.0%; Score 262.8; DB 24; Length 584;
Best Local Similarity 72.4%; Pred. No. 4.6e-39;
Matches 368; Conservative 0; Mismatches 137; Indels 3; Gaps 2;
QY 1215 GACAGTCCGGCG 1274
Db 507 GATAGTTCGGGTTACGCTTATTTAGGACGCGGATTTAGAAATACGATTACGTTAGTA 448
QY 1275 ACAGCGTTCATCTGACCTGGACGTGGCGACGAGTCTTCATCAAGCTGGACGCGGA 1334
Db 447 ATAGCGTATTTGTTATTTGGACGTGGCGACGAGGTTTATTAAGTTGGACGCGGA 388
QY 1335 AAGTCAGCGGCGCAACCAACAGTACAGACCTTCTCCGGCTTCATCATCTACCCG 1394
Db 387 AAGTCAGCGGCGTAATTAATAAGTATAGTATTTTTCGG-TTTATTTATTTTCG 329
QY 1395 ACTGAGCGGCG 1452
Db 328 ATGAGTCGGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCGTTTCG 269
QY 1453 TCTCTCCG 1512
Db 268 TTTTGTGCTGTTTATTCGAGCGGTTATTTTATTTTGGAGAGTTTCGCGGTGGAT 209
QY 1513 CTTCCGTTCCGGAGCGCGCGCTAAATGGCGAATCTTGTGCTCAAGGATTAAGTGG 1572
Db 208 TTTTTCGTTTCGTTGGCGGTTAAATGGCGAATTTTGTGCTTTAAGGTTAAGTGG 149
QY 1573 CCGGGAAGGAGGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1632
Db 148 TCGGGAAGGAGGAGGAGATTTCGGTTAGAGGAGTAGAGCGGATTTTCGGAGGATTTTCGTA 89
QY 1633 CCCAAGTCGCGCTGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1692
Db 88 TTTAAGTCGCGCTGGATTTATAGGGTAGAGGTCGTGCTTTTTCGTTTTCGTTATAGAGA 29
QY 1693 TGGGAGCAGTTTAAATAGCGGACTCA 1720
Db 28 TGGGAGTAGTTTAAATAGCGGATTA 1
RESULT 12
ID ABK10666
XX ABK10666 standard; cDNA; 981 BP.
XX
XX AC ABK10666;
XX
XX 05-JUN-2002 (first entry)
XX

CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a

PA (EPIG-) EPIGENOMICS AG.

XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX WPI; 2002-371829/40.
XX
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA -
XX
XX
PS ClaIm 12; 56pp + Sequence Listing; 56pp; German.
XX
XX This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-Cpg-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQ13410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
SQ Sequence 505 BP; 79 A; 45 C; 123 G; 258 T; 0 other;
Query Match 5.0%; Score 101.4; DB 24; Length 505;
Best Local Similarity 72.1%; Pred. No. 1.2e-09;
Matches 132; Conservative 0; Mismatches 51; Indels 0; Gaps 0;
QY 1218 AGGTCCGGCCAGCCGATTTGCTCAGGACGCGGACCAACTACGACTACGCCAGCAACA 1277
DB 65 AGGTCCGGGTTAGTGTATTGTTTATAGGACGCGGATTAGAAATACGATTACGTTAGTAATA 124
QY 1278 GCGTCATCTGACCTGGACGTGGCGGACGAGGCTCTTCATCAAGCTGGACGCGGGAAG 1337
DB 125 GCGTCATCTTGTATTGGACGTCGCGGACGAGGTTTTTATTAAAGTTGGATGGAGGTAAG 184
QY 1338 TGCACGGCGGCAACACCAAGTACAGCACCTCTCCGGCTTCATCATCTACCCGACT 1397
DB 185 TATACGGCGGTAAATAGTAATAATATAGTACGTTTTTTGGTTTTTATTATTTCGATT 244
QY 1398 GAG 1400
DB 245 GAG 247
Search completed: June 6, 2003, 18:36:34
Job time : 1344 secs

Result No.	Score	Query	Length	DB	ID	Description
1	86.2	4.3	1338	4	US-09-336-536-1	Sequence 1, Appli
2	85	4.2	1347	4	US-09-140-804-1	Sequence 1, Appli
3	83.8	4.2	728	4	US-09-336-536-2	Sequence 2, Appli
4	76.2	3.8	729	4	US-09-140-804-10	Sequence 10, Appli
5	73.6	3.7	2769	4	US-09-118-408-1	Sequence 1, Appli
6	73.6	3.7	2769	4	US-09-506-855-1	Sequence 1, Appli
7	67.2	3.2	4411529	4	US-09-103-840A-1	Sequence 1, Appli
8	64	3.2	4403765	4	US-09-103-840A-2	Sequence 2, Appli
9	62.2	3.1	1313	2	US-08-463-911-6	Sequence 6, Appli
10	62.2	3.1	4517	4	US-09-140-804-9	Sequence 9, Appli
11	61.8	3.1	4257	2	US-08-690-473-1	Sequence 1, Appli
12	61.8	3.1	4257	4	US-09-259-821A-1	Sequence 1, Appli
13	61.8	3.1	4257	4	US-08-843-659-1	Sequence 1, Appli
14	61.8	3.1	12001	1	US-08-458-568A-11	Sequence 11, Appli
15	60.8	3.0	4403765	4	US-09-103-840A-2	Sequence 2, Appli
16	60.4	3.0	4411529	4	US-09-103-840A-1	Sequence 1, Appli
17	60.2	3.0	11220	4	US-09-105-537-32	Sequence 32, Appli
18	60.2	3.0	36778	4	US-09-105-537-5	Sequence 5, Appli
19	59.6	3.0	390	4	US-09-197-649-9	Sequence 7, Appli
20	59.2	2.9	1276	2	US-08-463-911-1	Sequence 1, Appli
21	58.8	2.9	833	2	US-08-403-852D-3	Sequence 3, Appli
22	58.8	2.9	833	3	US-08-510-646B-3	Sequence 3, Appli
23	58.8	2.9	833	4	US-09-231-818-3	Sequence 3, Appli
24	58.8	2.9	5392	2	US-08-403-852D-1	Sequence 1, Appli
25	58.8	2.9	5392	3	US-08-510-646B-1	Sequence 1, Appli
26	58.8	2.9	5392	4	US-09-231-818-1	Sequence 1, Appli
27	58.2	2.9	1167	4	US-09-411-687A-4	Sequence 4, Appli

Query Match	4.23;	Score 85;	DB 4;	Length 1347;
Best Local Similarity	49.1%;	Pred. No. 3.6e-09;		
Matches 368;	Conservative 0;	Mismatches 370;	Indels 11;	Gaps 5;
Qy	663	GC	GGCGCGCGTCGGAGTGC	GCATGCTGCTGCTGCTGCTGGCCATCCCGCTGCTGTGTCG 722
Db	178	GGGCTCCCGTGCAGGCGCTATGAGGCCACTCCCTCGTCTGCTCTTGGGCTTGGCGG 237		
Qy	723	ACAGTCCCGCGGGCCACGCGCACATACAGAGATGCTGGGTGCTGCTGCCGATGCTGTGCGAC 782		
Db	238	CCGCTCGCCCCCACTGGAGGACAAAGATCCCCAGCCCTCTGCCGGGGCACCCCGGCC 297		
Qy	783	CGCATGGGCGCCCTGTGGCCCTGGTCCCGACGGCGGCGCTGCTTCCGTGCGCCCTTCCCGC 842		
Db	298	TTCCAGGACGCGCGGGCC---ACCATGGCAGCCAGGGCTTCCCGGGCGCGGATGGCGCGC 354		
Qy	843	CAGCGCCCAAGGGAGGTGGCGCGCGGGAAGCAGGGCTCGCGGGGGCCCCCTGGAC 902		
Db	355	ACGGCCGGACGGCGCGCCCGGGGCTCGGGAGAGAAAGGCGAGGGCGGGAGCGCGGAC 414		
Qy	903	CACCAAGTCCAAAGAGGGCCCCCAGGAGAAACCCGGCAGGCCAGGGCCCCCGGGGCCCTCCG 962		

```

RESULT 3
US-09-336-536-2
; Sequence 2, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 728
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-336-536-2

```

Query Match	4.28;	Score 83.8;	DB 4;	Length 728;
Best Local Similarity	49.4%;	Pred. No. 5.6e-09;		
Matches 360;	Conservative 0;	Mismatches 357;	Indels 12;	Gaps 5;
QY	683	ATGTTGCTGCTGCTGCTGGTGGCCATCCCGCTGCTGGTGACACAGCTCCCGCGGGCCAGCG	742	
Db	1	ATGAGGCACTCCTTCGCTGCTGCTGCTGGGCTGGCGCGCTGGCCCCACATGGAC	60	
QY	743	CACATACAGATGCTGGTGGTGGCGCTGCGCACCGCATGGTGCGACCCGATGGCGCCCTGGCCCT	802	
Db	61	GACAAACAGATCCCGACGCTCTGCCGGGGCACCCCGGCCCTCCAGSCAGCGCGGGCCAC	120	
QY	803	GGTCCGACGGCGGCGCTGCTTCGGTGCCCGCCCTTCCCGCCAGCGGCCAAGGAGAGGTG	862	
Db	121	CATGGCAGC-----CAGGGCTTGGCGGGCCCGCATGGCCGCGAGCGGCCGACGGCGCGCC	176	
QY	863	GGCGGGCGGGAAGCAGGCGCTTCGGGGGGCCCGCTTGACACACAGGTCCAGAGAGGGCCC	922	

Db 177 CGGGGCTCCGGGAGAGAAAGCGAGGGCGGAGCGGGGACTGCGGGGACCTCGAGGGGAC 236
 QY 923 CCAGGAGAAACCCGGCAGCGAGCCCGCGGGCCCTCCCGGTCCAGGTCCGGGCGGGGTG 982
 Db 237 CCCGGCGCGAGGAGGGGAGCCCGGGGCCACCGGGCT-CCCGGGAGTGCTC 295
 QY 983 GCCCGCGTCCGGGTACGTGCGCTCGCATGCTTTTACGCGGGCCCTGCGGGGCCCCAC 1042
 Db 296 GGTGCTCCGCGATCGCCCTTCAGCGCAAGCGCTCCGAGAGC- ----CGGGTGCTCCG 350
 QY 1043 GAGGGTTACAGGTGCTGCGCTTCGACGAGCTGCTGACCAACGTGGGCAACGCTACGAG 1102
 Db 351 CGGTCTGAGCGCACCTTGCGCTTCGACCGGCTGCTGGTGAACGAGCAGGACATTACGAC 410
 QY 1103 GCAGCAGGGGCAAGTTTACTTGGCCCATCGCCAGGCGTCTACTTCTTCGCTTACCACTG 1162
 Db 411 GCCTCACCGGCAAGTTTCACTTGCAGAGTGCTTGGGGTCTACTACTTTCGCGCTCATGCC 470
 QY 1163 CTCATCGGGGCGGC-GAGCGGACACAGCATGTGGGCGGCGACCTCATGAAGAAGCGACAGGT 1221
 Db 471 ACCGTCTACCGGCGCACCTTGCAATTTGATCTGTTGAAGAATGGCGAATCCATTGCTCT 530
 QY 1222 CCGGGCAGCGCATTTGCTCAGACCGCGGACCAAGACTAGGACTAGCCAGCAACAGCGT 1281
 Db 531 TTCTTCAGGTTTTCGGGGGGTGGCCCAAGCCAGC-CTCGCTCGGGGGGGGCCATGGT 589
 QY 1282 CATTTGCACTGGAGCTGGCGGACGAGGTCTTTCATCAAGCTGGAGCGGGGAAAGTGCA 1341
 Db 590 GAGGCTGGAGCTGAGGACCAAGTGTGGGTGCAGGTGGGTGCTACTACATTGGCAT 649
 QY 1342 CGCGGCGACACCAAGTACAGCACCTTCTCCGGCTTCATCATCTACCCGAGTACG 1401
 Db 650 CTATGCGAGCATCAAGACAGACAGCACCTTCTCCGGATTCTGGTACTCCGACTGGCA 709
 QY 1402 CGGCGCCCGC 1410
 Db 710 CAGTCCCC 718
 RESULT 4
 US-09-140-804-10
 ; Sequence 10, Application US/09140804
 ; Patent No. 6197930
 ; GENERAL INFORMATION:
 ; APPLICANT: Sheppard, Paul O.
 ; APPLICANT: Humes, Jacqueline M.
 ; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
 ; FILE REFERENCE: 97-49
 ; CURRENT APPLICATION NUMBER: US/09/140,804
 ; CURRENT FILING DATE: 1998-08-26
 ; EARLIER APPLICATION NUMBER: 60/056,983
 ; EARLIER FILING DATE: 1997-08-26
 ; NUMBER OF SEQ ID NOS: 47
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 10
 ; LENGTH: 729
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Degenerate nucleotide sequence encoding the zsig39
 ; OTHER INFORMATION: polypeptide of SEQ ID NO:2.
 US-09-140-804-10
 Query Match 3.8%; Score 76.2; DB 4; Length 729;
 Best Local Similarity 29.8%; Pred. No. 2.1e-07;
 Matches 211; Conservative 80; Mismatches 412; Indels 9; Gaps 2;
 QY 690 TGCTGCTGTGTGGGCAATCCCGCTGCTGTGTGCACAGCTCCCGCGGGCGAGCGCACTACG 749
 Db 2 TGMGNCNTYNTNGTNTYNTYNTYNTNGGNTNCGNGGNGNWSNCCNCCNTYNGAYG 61
 QY 750 AGATGTGGTGCCTGCCCATGGTGTGCGACCCCGCATGGGCCCGCTGGCCCTGGTCCCG 809

[illegible]

CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: United States
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/843,659
FILING DATE: Concurrently Herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L.
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: ARSB:519
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 4257 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-843-659-1

Query Match	3.1%	Score 61.8	DB 4	Length 4257
Best Local Similarity	47.6%	Pred. No. 0.00032		
Matches 245	Conservative 0	Mismatches 267	Indels 3	Gaps 2
QY	467	CGCGCGCTGGAGCCGCGCAGAGGACAGGTAAGGGAGCGGGGACCGCTGCGGGAGT	526	
Db	2211	CGCGGACGGCCCGCGCTGGCGCGTGGCTGCGGAGCTGCGGTTCTGGCGACCGCT	2270	
QY	527	GCAGACCCAGGCCCAAGCGGGTACCGCTCTCTGGCCCCCGG--AGAGCCCGCGGCCCGCG	585	
Db	2271	GGTGCTCATGCGCTTCGCGGGGACCTGCGCGTGGCGCGGCGGACGAGCGCGCTGGC	2330	
QY	586	AGCCATTGGCCCCAAGATGATGGAAGATTGCTGGGCCCTGGCAGCGTTCGGGGTGAACCG	645	
Db	2331	CGCGCTGCGCGCGTGAAGCTTGGTCCCGGGGCCCTGGCGCCGCGCTGCGCGGACCC	2390	
QY	646	GCSCAAGAGGTGGCGGGCGCGCGCTCGGAGTGGCCATGGTGTCTCTGCTGTGTGGC	705	
Db	2391	GCBCCTGCGAGCTCCGGGCGCCCGCGCGCGGACCTGCTGTTTGACAAACAGAGCCT	2450	
QY	706	CATCCCGCTGTGGTGCACAGCTCCCGCGGGCGCAGCGCACTACGAGATGTG--GGTCGC	763	
Db	2451	GCGCCCTCTGCTGGCGGCGCGGCACGCGACCGGACGCGCGCGCGCGCTGGCGCGC	2510	
QY	764	TGCGCGCATGTTGCGACCGCGCATNGGCCCCCTGGGCCCTGGTCCCAGCGGGCGCTGT	823	
Db	2511	CGCGCTCTCGCGCGCGCGGAGGGGGGCAAGCGCAAGAGTCCCGGGCGCGCGCGCGC	2570	
QY	824	TCGCTGCCCCCTTCCCGCAGCGCGCAAGGAGAGTGGCGCGCGCGGGAAGAGCGC	883	
Db	2571	GCCTGGAGGGCGGGCGCGCGACCCCGAAGACGAAGAGGGGGCGGACGCCCGCGG	2630	
QY	884	CTCGGGGGGCCCCCTTGACACACAGTCCAAAGAGGGCCCCAGGAGAACCGCAGCGCA	943	
Db	2631	CTCGAGCGCGCGCCCCCTTCCCGCGCGCGCGCCCCCTCCACGCCCGGGGCCCCGA	2690	
QY	944	GGCCCCCGCGGCCCTTCCCGGTCACAGTTCGCGGGCGG	978	
Db	2691	GCCTCGCCCGCGCGACCGCGCGCGCGCGCGCGCGG	2725	

RESULT 14

```

US-08-458-568A-11/c
; Sequence 11, Application US/08458568A
; Patent No. 5821339
; GENERAL INFORMATION:
; APPLICANT: Schaffer, Priscilla A.
; APPLICANT: Yeh, Lily
; TITLE OF INVENTION: Compositions and Methods for Treatment of Herpesvirus
; TITLE OF INVENTION: Infections
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock, Washburn, Kurtz, Mackiewicz & No. 5821339-r1s
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/458,568A
; FILING DATE: 02-JUNE-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/065,146
; FILING DATE: 05-MAY-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Leary Ph.D., Kathryn R.
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: DFCI-0029
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100
; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12001 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Herpes simplex virus
; STRAIN: Herpes Simplex Virus Type 1
; US-08-458-568A-11

Query Match 3.1%; Score 61.8; DB 1; Length 12001;
Best Local Similarity 47.6%; Pred. No. 0.0004;
Matches 245; Conservative 0; Mismatches 267; Indels 3; Gaps 2;

QY 467 CGCGCGCCTGGAGCCCGCCACAGGACGAGGTAAGGGAGCGGGGCGACCCGCTCGGGGGAGT 526
      |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 4463 CGCGAGCGCCCGCGCGCTGCGCGCGTGGCTGCGCGAGCTCGGTGTCGCGCAGCGCGCT 4404
      |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
QY 527 GCAGACCGAGCGCCCAAGCGGGGTCACCGCTTCCTGGCCCGCGG-AGAGCCCGCGCCCGCGC 585
      |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 4403 GGTGCTCATCGCCTGCGCGGGGACCTGCGCGTGGCGGGCGGAGGAGCCCGCTGGC 4344
      |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
QY 586 AGCATTTGCCCAAGAGTAGGAAAGATTTCGTGGCCCTGGCAGCGTTCGGCGCTGAGCCG 645
      |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 4343 CGCGCTGCGCGCGGTGAGCGCTGTCGCGGGGGCCCTGGGCGCGCGCTGCGCGGGACCC 4284
      |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
QY 646 GCACAAGAGGTTGGCGGCGCGCGCTGCGAGTGGCCATGCTGCTGCTGTGTGTGGC 705
      |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 4283 CGCGCTGCCAGTCTCCGCGCGCGCGCGCGCGGACCTGCTGTTTGACAACACAGAGCT 4224
      |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
QY 706 CATCCCGCTGCTGTCACAGTCCCGCGGGCGAGCGCACTACAGATGCTG--GGTCGC 763
      |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||
DB 4223 CGCGCCCGCTGCTGCGCGCGCGCGCAAGCGCACCGGACCGCGACGCGCTGGCGCGCGC 4164
      |||||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||  ||

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 20:04:08 ; Search time 308 Seconds

(without alignments)
9160.608 Million cell updates/sec

Title: US-10-005-499-377

Perfect score: 2016

Sequence: 1 aggaagtttagggaggcgga.....tggggctacagctgctg 2016

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 870385 seqs, -699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/ptodata/2/pubpna/PCT05_PUBCOMB.seq:*

7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*

10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*

12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*

13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	269	13.3	466	10	US-09-867-701-1219
2	258.8	12.8	795	9	US-09-998-563-1
3	237.8	11.8	750	9	US-09-998-563-3
4	181.4	9.0	804	9	US-09-998-582-1
5	176.4	8.8	804	9	US-09-998-582-3
6	98.2	4.9	1136	10	US-09-917-800A-1671
7	95.2	4.7	15558	9	US-09-764-872-823
8	93.2	4.6	600	10	US-09-864-761-11941
9	91	4.5	1185	9	US-10-098-841-182
10	88.2	4.4	1377	12	US-10-044-090-60
11	86.2	4.3	1338	9	US-09-796-753-67
12	85	4.2	1377	9	US-09-944-413-41
13	85	4.2	1377	9	US-09-944-403-41
14	85	4.2	1377	9	US-09-944-896-41
15	85	4.2	1377	9	US-09-944-941-41
16	85	4.2	1377	9	US-09-944-907-41
17	85	4.2	1377	9	US-09-944-929-41
18	85	4.2	1377	9	US-10-028-072-361
19	85	4.2	1377	9	US-10-121-049-361

20 85 4.2 1377 9 US-10-123-904-361 Sequence 361, App

21 85 4.2 1377 9 US-10-140-470-361 Sequence 361, App

22 85 4.2 1377 9 US-10-175-746-361 Sequence 361, App

23 85 4.2 1377 9 US-10-176-918-361 Sequence 361, App

24 85 4.2 1377 9 US-10-176-921-361 Sequence 361, App

25 85 4.2 1377 9 US-10-137-865-361 Sequence 361, App

26 85 4.2 1377 9 US-10-140-474-361 Sequence 361, App

27 85 4.2 1377 9 US-10-142-431-361 Sequence 361, App

28 85 4.2 1377 9 US-10-143-114-361 Sequence 361, App

29 85 4.2 1377 9 US-10-140-002-361 Sequence 361, App

30 85 4.2 1377 9 US-10-142-419-361 Sequence 361, App

31 85 4.2 1377 9 US-10-123-262-361 Sequence 361, App

32 85 4.2 1377 9 US-10-142-423-361 Sequence 361, App

33 85 4.2 1377 9 US-10-121-050-361 Sequence 361, App

34 85 4.2 1377 9 US-10-141-755-361 Sequence 361, App

35 85 4.2 1377 9 US-10-143-032-361 Sequence 361, App

36 85 4.2 1377 9 US-10-123-108-361 Sequence 361, App

37 85 4.2 1377 9 US-10-123-236-361 Sequence 361, App

38 85 4.2 1377 9 US-10-123-261-361 Sequence 361, App

39 85 4.2 1377 9 US-10-140-921-361 Sequence 361, App

40 85 4.2 1377 9 US-10-140-928-361 Sequence 361, App

41 85 4.2 1377 9 US-10-121-045-361 Sequence 361, App

42 85 4.2 1377 9 US-10-123-292-361 Sequence 361, App

43 85 4.2 1377 9 US-10-123-903-361 Sequence 361, App

44 85 4.2 1377 9 US-10-124-819-361 Sequence 361, App

45 85 4.2 1377 9 US-10-124-822-361 Sequence 361, App

ALIGNMENTS

RESULT 1

US-09-867-701-1219

; Sequence 1219, Application US/09867701

; Patent No. US20020132237A1

; GENERAL INFORMATION:

; APPLICANT: Aglate, Paul A.

; APPLICANT: Jones, Robert

; APPLICANT: Harlocker, Susan L.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.497

; CURRENT APPLICATION NUMBER: US/09/867,701

; CURRENT FILING DATE: 2001-05-29

; NUMBER OF SEQ ID NOS: 10912

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1219

; LENGTH: 466

; TYPE: DNA

; ORGANISM: Homo sapien

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)....(466)

; OTHER INFORMATION: n = A,T,C or G

US-09-867-701-1219

Query Match 13.3%; Score 269; DB 10; Length 466;

Best Local Similarity 96.8%; Pred No. 6.8e-61;

Matches 338; Conservative 0; Mismatches 5; Indels 6; Gaps 6;

QY 1606 GAGCGACTTCGGAGGGATACCCGACCCAGTCGCGCTGGAGCCCATAGGGCAGAG 1665

Db 1 GAGCGACTTCGGGA-GGATACCCGACCCCAAGTG-GCGCTGGACCCCATATA-GGGCAGAG 57

QY 1666 GTCGTGGCTTCTCTTTTGTACAGAGATGGGGAGGAGTTTAAATAGCGGGACTCAGAGC 1725

Db 58 GTCGTGGCTTCTCTTTTGTACAGAGATGGGGAGGAGTTTAAATAGCGGGACTCAGAGC 117

QY 1726 CCAGAAAGCCGGAGGAGCCCGCAGCTTTCGAGGGGAAATAACAGAAACAGAGGAGAGC 1785

Db 118 CCAGAAAGCCGGAGGAGCCCGCAGCTTTCGAGGGGAAATAACAGAAACAGAGGAGAGC 177

QY 1786 CCATTTAGGCAAGAGAGACATTAACACAGGGTAGTGCAGGTTTCTCCGTCACAACTTTCT 1845

Db 178 CCATTAGCAAGAGAACATTAACAGGAGGTAGTCAGGTTCTCCGTCACAACCTTCT 237
QY 1846 CTCGCCACCTCTCGTCCCTCGTCTCCACTTTTCAGGCTCAGGCTCCAGCCTTGCGAGCC 1905
Db 238 CTCGCCACCTCTCGTCCCTCGTCTCCACTTTTCAGGCTCAGGCTCCAG-CTTGGCAGCC 296
QY 1906 TTCCTGTGAACCT-GGAGGAACCACT-GAATTCCTTCTCCCTGGCATTTTAAAA 1952
Db 297 TTCCTGTGAACCTGGGAGGAACCACTGGAATTCCTTCTCCCTGGCATTTTAA 345
RESULT 2
US-09-998-563-1
; Sequence 1, Application US/09998563
; Patent No. US20020155546A1
; GENERAL INFORMATION:
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: ADIPOCYTE COMPLEMENT RELATED PROTEIN
; FILE REFERENCE: ZACRP12
; CURRENT APPLICATION NUMBER: US/09/998,563
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 60/253,867
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 795
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (46)...(795)
US-09-998-563-1

Query Match 12.8%; Score 258.8; DB 9; Length 795;
Best Local Similarity 70.9%; Pred. No. 3.4e-58;
Matches 343; Conservative 0; Mismatches 141; Indels 0; Gaps 0;
QY 912 CAAGAGGGCCCCAGGAGAACCCGGCAGCCAGGCCCCCGGGCCCTCCGGTCCAGGTC 971
Db 311 CCAGGGGCGTCCGGGTGGTGGCGCGCGGGCGGGGTAGTGGCGATTCCGAGGGTGAAG 370
QY 972 CGGGGGGGTGGCGCCCTGCGGCTACGTCGCTCGCATTCCTTCTACGGCGGCTGC 1031
Db 371 TGACCACTGCGTACGGCGCCACCTTCAGCGGCCCAAGATCGCTTCTATGTGGTCTCA 430
QY 1032 GGGGGCCCCACAGAGGGTTACGAGGTGCTGCGCTTCGACGACGTGTGACCAACGTGGGA 1091
Db 431 AGAGCCCCCAGCAAGGCTATGAGGTGCTGAAGTTCGATGACGTGGTCAACCACTCGGCA 490
QY 1092 ACGCTACAGGACGCCAGCGCAAGTTTACTTGCCCATGCGCAGCGCTCTACTTCTCG 1151
Db 491 ATCACTATACCCACCGGCAAGTTTCAGCTGCGCAGGTACGCGGCATCTACTTCTCA 550
QY 1152 CTTACACCTGCTCATCGCGCGCGCGCAGCCAGCATGTGGCGCGACCTTCATGAAGA 1211
Db 551 CTTACACATCTCATGCGCGCGCGCGGCGGACCATGTGGCGGACCTCTCAAGA 610
QY 1212 ACGGACAGTCCGGCGCAGCGCCATGTCTCAGGACGGGACCAAGCACTACGACTACGCA 1271
Db 611 ACGGCGAGGTCCGGCGCAGCGCATTTGACAGAGGCGCGCAGCACTACGACTACGCA 670
QY 1272 GCAACAGCTCATCTGCACTGAGCGTGGGCGGAGAGGTCTTCATCAAGCTGGAGCGG 1331
Db 671 GTACAGCGTGTGCTGCACTTGGATTACGGGAGCAAGGTATGTGAGAGCTGGATGGCG 730
QY 1332 GGAAAGTCCAGCGCGCAACCAACCAAGTACAGCACCTTCTCCGGCTTCATCATCTACCC 1391
Db 731 GGAAGGCTCAGCGAGGCAATAACAAGTACAGCACGTTCTCTCGGCTTCTCTCTGAC 790
QY 1392 CCGA 1395

Db 791 CGGA 794
RESULT 3
US-09-998-563-3
; Sequence 3, Application US/09998563
; Patent No. US20020155546A1
; GENERAL INFORMATION:
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: ADIPOCYTE COMPLEMENT RELATED PROTEIN
; FILE REFERENCE: ZACRP12
; CURRENT APPLICATION NUMBER: US/09/998,563
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 60/253,867
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate polynucleotide encoding a polypeptide
; OTHER INFORMATION: of SEQ ID NO:2
; NAME/KEY: variation
; LOCATION: (1)...(750)
; OTHER INFORMATION: Each n is independently T, A, G, or C
; NAME/KEY: misc_feature
; LOCATION: (1)...(750)
; OTHER INFORMATION: n = A,T,C or G
US-09-998-563-3

Query Match 11.8%; Score 237.8; DB 9; Length 750;
Best Local Similarity 52.0%; Pred. No. 1e-52;
Matches 213; Conservative 85; Mismatches 112; Indels 0; Gaps 0;
QY 987 CGCTCGCGGCTACGCTCGCTTCCTAGCGGGCTGCGGGCGGCCCGCCAGGAG 1046
Db 341 SNGCNACNTTYSNGNCCNAAATATGCTTTTATGTGNTGNTNAAWNSNCCNAYGARG 400
QY 1047 GTTACGAGTGTGCTGCTTCGACGAGTGTGTCACCAACCTGGGCAACGCTACGAGGAG 1106
Db 401 GNTAYGARTNTNAARTTYGAYGTGNTNACNAAYTNGNNAAYTAYGAYCCNA 460
QY 1107 CCAGCGGCAAGTTTACTTCCCGCATGCCAGGCGTCTACTTCTGCTTACCAGCTCTCA 1166
Db 461 CNACNGNNAARTTYWSNTGYCARGTNGMNGGNATHTATYTTTAYACNTAYATHTYNA 520
QY 1167 TCGCGGGCGGCGCAGCCAGCATGTGGCGGAGCTCATGAAGAACGACAGGTCGGG 1226
Db 521 TGMNGNGNGNGAYGNACNWSNATGTGGCGNGAYTNTGYAARAYGNCARGTNMNG 580
QY 1227 CCAGCGCATCTGCTCAGGACGGGACCAAGTACGACTACGCCAGCAACAGCGCTATTC 1286
Db 581 CNWSNGCNATHGNCARGAYGNGAYCARAAATAYGAYTAYGCMNSNAAYWSNGTNGTY 640
QY 1287 TGCACCTGACGTGGCGGACGAGGTCTTCATCAAGCTGGACGCGGAAAGTGCACGGG 1346
Db 641 TNCATYTTNGAYWSNGNGAYGARGTNTATGTNARTNGAYGNGNGNNAARGNCATYGG 700
QY 1347 GCAACACCAACAAAGTACAGCACCTTCTCCGGCTTCATCATCTACCCCGAC 1396
Db 701 GNAAYAAAYAAARTYWSNACNTTYSNGGNTTYYTNTYNTAYCNGAY 750

RESULT 4
US-09-998-582-1
; Sequence 1, Application US/09998582
; Patent No. US20020160474A1
; GENERAL INFORMATION:
; APPLICANT: Fox, Brian

US-09-998-582-3

Db 553 GTNTNCAYTNGARCCNGGNGAYGARGTNTAYATHAARTYNGAYGGNGNAAARGNCAY 612

Qy 1343 GCGCGCACACCAACAAGTACAGCACCTTCTCCGGCTTCATCATC 1387

Db 613 GGNGNAAAYAAAYAAATAYWSNACNTTYTNGAYVNTNTNTTY 657

RESULT 6
US-09-917-800A-1671
Sequence 1671, Application US/09917800A
Patent No. US20020119462A1
GENERAL INFORMATION:
APPLICANT: Mendrick, Donna
APPLICANT: Porter, Mark
APPLICANT: Johnson, Kory
APPLICANT: Castle, Arthur
APPLICANT: Elashoff, Michael
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Molecular Toxicology Modelling
FILE REFERENCE: 44921-5038-US
CURRENT APPLICATION NUMBER: US/09/917,800A
CURRENT FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: US 60/222,040
PRIOR FILING DATE: 2000-07-31
PRIOR APPLICATION NUMBER: US 60/222,880
PRIOR FILING DATE: 2000-11-02
PRIOR APPLICATION NUMBER: US 60/290,029
PRIOR FILING DATE: 2001-05-11
PRIOR APPLICATION NUMBER: US 60/290,645
PRIOR FILING DATE: 2001-05-15
PRIOR APPLICATION NUMBER: US 60/292,336
PRIOR FILING DATE: 2001-05-22
PRIOR APPLICATION NUMBER: US 60/295,798
PRIOR FILING DATE: 2001-06-06
PRIOR APPLICATION NUMBER: US 60/297,457
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,884
PRIOR FILING DATE: 2001-06-19
PRIOR APPLICATION NUMBER: US 60/303,459
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 1740
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1671
LENGTH: 1136
TYPE: DNA
ORGANISM: Rattus norvegicus
FEATURE:
OTHER INFORMATION: Genbank Accession No. US2002020119462A1

[illegible]

508	GA	AAAGGTGATCACC	AATGTTAATGATAC	TACTACGACCGCGG	CAGTGGCAGTTCA	CTGC	667
1127	CC	CATGCCAGGCGCT	ACTCTTCTTCGTT	TACCACGTGCT	CATGCGCGCGCGG	CACGAC	1186
668	AA	GTACCTGGCCT	CTACTACTTCA	CTACCCAGCCAGT	TCCGCGGGAATCT	GTGTG	727
1187	AG	CATGTGGCGCC	GACCTCATGA	AGACGGACAGGT	CCGGGCCAGCCCA	TGCTCAGGAC	1246
728	AA	CATCGTGGCG	CGCGCAGACCG	CATGCAGAAAGT	TCTCACTTCTCG	CACTAT	787
1247	GC	GACCAACAACT	ACGACTACGCC	AGCAGACGGT	CATTCTGCACCT	TGGACGTGGCGG	1306
788	GC	CAAAACACCT	TCCAGTTCAC	CGGGTGGGTAGT	CTTTGAAGCTG	GACGACGGAAG	847
1307	GAG	GTCTTCAT	CAAGCTGGAC	GCGGGAAGTGC	ACGGGCAACACCA	CAAGTACAGC	1366
848	GT	TGTTCACT	TGCAGGCCA	CAGACAAGACT	CTCCTCGTGGCG	CTCAGGGAGCCAT	907
1367	AC	CTTCTCCGGCT	TTCATCATCT	ACCCCGATGAG	CGGCGCCCGCCCG	CTGCCCGCTCG	1426
908	AT	CTTCACTGGCT	TTCTGCTTTT	CCCTGACATGAT	GTATGATACAGGG	GTCAATCACT	967
1427	CC	CTTCTCTCC	GTGCTTACC	CACTCTCTG	CCCGCCCGAC	CCGAGCGCGCAC	1481
968	CC	TATCAAAACCT	CTCTCCCTG	CAGTAATCTCT	CCCTGGAACCC	CCAGACACTGGCC	1022

```

RESULT 7
US-09-764-872-823/c
; Sequence 823, Application US/09764872
; Publication No. US20030050231A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P125
; CURRENT APPLICATION NUMBER: US/09/764,872
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 957
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 823
; LENGTH: 15558
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-872-823

```

Query Match	4.7%	Score 95.2;	DB 9;	Length 15558;
Best Local Similarity	73.4%	Pred. No. 3.1e-15;		
Matches 135;	Conservative 0;	Mismatches 48;	Indels 1;	Gaps 1;
Qy	1216	ACAGGTCGGGGCAGCGCCATTGCTCAGGACGGGGACCAAGCTACGACTACGCCAGCAA	1275	
Db	9700	ATAGGTCGGTCTAGTGCATTTCCCNAGATGCTGATCAGATTACGACTATGCCAGTAA	9641	
Qy	1276	CAGCGCTCATTTGCACCTTGGACGTGGCGGACGAGGTCTTTCATATCAAGCTGGACGGCGGGAA	1335	
Db	9640	CAGTGTGGTTCTTCAATTTGGAGCGGGAGATGAAGTCTATATCAAAATTAGATGGCGGGAA	9581	
Qy	1336	AGTGCAGCGGGCAACCAACCAAGTACAGCACCTTCTCGGGCTTCATCATCTACCCCGA	1395	
Db	9580	AGCCATGGAGGAACAAACAACAANTACAGCAGCTT-CTGGATTATTTATTATCTCTGA	9522	
Qy	1396	CTGA	1399	
Db	9521	CTGA	9518	

RESULT 8
US-09-864-761-11941
; Sequence 11941, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.

[illegible]


```

Db      444  TGCCTGGAGCTCGAGGGAGCCCGGCGCCGAGGAGAGCGGGAGCCCGGGGGCCACCG 503
QY      963  GTCCAGGTCGGGGCGGGTGGCGCGCTGCGCGGTACGCGCTCGCATTTCTTCTAGG 1022
Db      504  GGCC-TGCCGGGGAGTGTCTGGTGGCTCCGCGATCCGCTTCAGCGCCAAAGCGCTCCGAG 562
QY      1023  CGGGCCTGCGGGGGCCCGCCACGAGGTTACGAGGTGCTGGGCTTCGACGACGTGGTGACCA 1082
Db      563  AGC-----CGGTGGCTCGCGCGTCTGACGCAACCTTGCCCTTCGACCGCGTGTGGTGA 617
QY      1083  ACGTGGGCAACGCCCTACGAGGAGCGAGCGGCAAGTTTACTTGGCCCATGCGCAGGCGTCT 1142
Db      618  ACGAGCAGGGACATTACGACGCGCTCACCGGCAAGTTACCTGCCAGGTGCGCTGGGGTCT 677
QY      1143  ACTTCTTCGCTTACCACGTGCTCATGCGCGCGCGGC-GACGGCACCAGCATGTGGCCCGAC 1201
Db      678  ACTACTTCGCCCTCCATGCCACCGTCTACCGGGCCAGCGCTGCGATTGATCTGGTGAAGA 737
QY      1202  CTCATGAAGAAGCGGAGGTCCGGGCGCAGCGCATTTGCTCAGGAGCGCGGACCAAGAACTAC 1261
Db      738  ATGGCGAATCCATTGCTCTTCTTCCAGTTTTCGGGGGGTGGCCCAAGCCAGC-CTCG 796
QY      1262  GACTAGCGCCAGCAACAGCGTCAATCTGCACTGGAGCGTGGGGCGAGAGGTCTTCATCAAG 1321
Db      797  CTCTCGGGGGGGCCATGTTGAGGCTGGAGCTGGAGCCAAAGTGTGGTGCAGGTGGGT 856
QY      1322  CTGGAGCGGGGAAAGTGCACGCGCGCAACACCAAGTACAGCAGCCTTCTCCGGCTTC 1381
Db      857  GTGGGTGACTACATTGGGCATCTATGCCAGCATCAAGACAGACAGCAGCCTTCCTCCGATT 916
QY      1382  ATCATCTACCCGAGTGGCGGCCCGCC 1410
Db      917  CTGGTGTACTCCGACTGGCACAGCTCCCC 945
```

Search completed: June 7, 2003, 01:08:04
Job time : 315 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 6, 2003, 17:46:03 ; Search time 2931 Seconds
(without alignments)
11139.577 Million cell updates/sec

Title: US-10-005-499-377

Perfect score: 2016

Sequence: 1 aggaaggttagggagcgga.....tggggctacagctgctgctg 2016

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:**

2: em_esthum:**

3: em_estin:**

4: em_estmu:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_htc:**

9: gb_est1:**

10: gb_est2:**

11: gb_htc:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: gb_gss:**

18: em_gss_hum:**

19: em_gss_inv:**

20: em_gss_pln:**

21: em_gss_vrt:**

22: em_gss_fun:**

23: em_gss_mam:**

24: em_gss_mus:**

25: em_gss_Other:**

26: em_gss_pro:**

27: em_gss_rod:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	958	47.5	967	9	AL561826
c 2	874.8	43.4	914	9	AL529897
c 3	869.8	43.1	920	9	AL519829
c 4	868.6	43.1	942	9	AL529898
c 5	843.2	41.8	894	9	AL518440
c 6	824.8	40.9	870	9	AL518439

7	479.8	23.8	557	14	BM762336	BM762336 K-EST0043
8	355.8	17.6	387	12	BE888168	BE888168 601511679
9	353.8	17.5	425	14	BM844470	BM844470 K-EST0122
10	327.8	16.3	845	13	BI603927	BI603927 603244539
11	314.4	15.6	542	13	BM141973	BM141973 1f25a05.y
12	314.4	15.6	552	13	BI793113	BI793113 1e48a10.y
c 13	313.4	15.5	323	17	AQ200834	AQ200834 RPCI11-62
14	308.8	15.3	461	13	BI793031	BI793031 1e47a04.y
15	307.6	15.3	540	13	BM354017	BM354017 1f29d01.y
16	304.2	15.1	638	12	BF686380	BF686380 602143757
17	297.8	14.8	301	14	BM847021	BM847021 K-EST0126
c 18	295.2	14.6	389	17	AZ793063	AZ793063 2M0045F17
19	288.2	14.3	982	12	EG169143	EG169143 602320603
c 20	286.2	14.2	307	17	AQ586940	AQ586940 RPCI-11-4
c 21	269.8	13.3	466	14	BM128150	BM128150 1f10d01.y
22	268.6	13.3	533	14	H61343	H61343 yu39g04.r1
23	268.6	13.3	533	14	H6101126	H6101126 1j24c07.y
24	268.6	13.3	578	14	EQ101250	EQ101250 1j25h09.y
25	266	13.0	535	14	EQ101230	EQ101230 1j25f09.y
26	261.4	13.0	644	14	BM711545	BM711545 UI-E-CQ1-
27	259.4	12.9	491	13	BM128991	BM128991 1f16b02.y
28	257.4	12.8	523	13	BM128566	BM128566 1f13a01.y
c 29	256.2	12.7	654	13	BJ092454	BJ092454 BJ092454
c 30	255.6	12.7	334	12	BF593076	BF593076 na331h12
31	255.4	12.7	531	12	EG294846	EG294846 602391671
c 32	240.6	11.9	885	14	EQ171748	EQ171748 AGENCOURT
c 33	236	11.7	782	9	AL567669	AL567669 AL567669
34	230	11.4	929	13	BI737761	BI737761 603357310
35	221	11.0	842	12	BG325001	BG325001 602423434
c 36	219.8	10.9	815	9	AL566412	AL566412 AL566412
37	219.4	10.9	297	10	BE650685	BE650685 UI-M-BH3-
38	217	10.8	815	12	BF786942	BF786942 602112135
c 39	198.2	9.8	518	12	BE944830	BE944830 UI-M-BH3-
40	197.2	9.8	518	12	BF471505	BF471505 UI-M-BH3-
41	188.6	9.4	383	13	BI104863	BI104863 602892640
c 42	188.4	9.3	335	13	BI042421	BI042421 RC5-OT009
43	186.4	9.2	318	13	BI042438	BI042438 RC5-OT009
c 44	184.2	9.1	487	12	BG374818	BG374818 UI-R-CV1-
c 45	181.6	9.0	391	9	AI683209	AI683209 tx02a08.x

ALIGNMENTS

RESULT 1
LOCUS AL561826/c
DEFINITION AL561826 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB005YK05 3 prime, mRNA sequence.
ACCESSION AL561826
VERSION AL561826.1 GI:12909642
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 967)
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1..967
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB005YK05"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"


```
Db 724 AACTACGACTACGACGACGACGCTCATCTGACCTGGAGTGGCGAGAGGTCTTC 665
Qy 1316 ATCAAGCTGGACGGGGGAAAGTACAGCGGGGCAACACCAACAAAGTACAGACACTTCTCC 1375
Db 664 ATCAAGCTGGACGGGGGAAAGTACAGCGGGGCAACACCAACAAAGTACAGACACTTCTCC 605
Qy 1376 GGCTTCATCATCTACCCGACTGAGCGGGGCCCCCGCGCTGCCCCCGCTGCCCTTCTTC 1435
Db 604 GGCTTCATCATCTACCCGACTGAGCGGGGCCCCCGCGCTGCCCCCGCTGCCCTTCTTC 545
Qy 1436 TCCGCTCTCACCACACTCTCTGCGGGGCCCCCGCGCTGCCCCCGCTGCCCTTCTTC 1495
Db 544 TCCGCTCTCACCACACTCTCTGCGGGGCCCCCGCGCTGCCCCCGCTGCCCTTCTTC 485
Qy 1496 CTGCGGCTGGGGTGGAGCCCTTCCGCTTCCGAGGCGGCTAAATGGGGAACCTTCTGGTG 1555
Db 484 CTGCGGCTGGGGTGGAGCCCTTCCGCTTCCGAGGCGGCTAAATGGGGAACCTTCTGGTG 425
Qy 1556 CTCAAGGGTATAAGTGGCGGGGAAGAGAGAGAGACCCGCGGCAAGAGAGAGAGACTTC 1615
Db 424 CTCAAGGGTATAAGTGGCGGGGAAGAGAGAGAGAGACCCGCTCCAGRGAGAGAGACTTC 365
Qy 1616 CGGAGGATCACCCGACCCCAAGTCCGCTGGACCCCATAGGGGCGAGAGGTCTGGCTT 1675
Db 364 CGGAGGATCACCCGACCCCAAGTCCGCTGGACCCCATAGGGGCGAGAGGTCTGGCTT 305
Qy 1676 TCTCTTTGTACAGATGGGAGCAGTCTTAAATAGCGGACTCAGAGGCGGCAAGAGCC 1735
Db 304 TCTCTTTGTACAGATGGGAGCAGTCTTAAATAGCGGACTCAGAGGCGGCAAGAGCC 245
Qy 1736 GGAGGGAAGCCCCGCGAGCTTGGCGAGGAAATACAGAAACAGGAGGAGCCATTTAGCC 1795
Db 244 GGAGGGAAGCCCCGCGAGCTTGGCGAGGAAATACAGAAACAGGAGGAGCCATTTAGCC 185
Qy 1796 AAGAGAGACATTAACACAGGAGTGTAGTTCGCTGCTCAACTTCTCTGCCAGCC 1855
Db 184 AAGAGAGACATTAACACAGGAGTGTAGTTCGCTGCTCAACTTCTCTGCCAGCC 125
Qy 1856 TCTCGTCCCTCGCTCCACTTTCAGGCTCAGGCTCAGGCTCGGAGGCTTCTCTGTGAA 1915
Db 124 TCAGTCCCGWACTTCACTTTCAGGCTCAGGCTCAGGCTCGGAGGCTTCTCTGTGAA 65
Qy 1916 CTGAGGAACAGTGAATTTCTTCTGCGATTTAAACGCAATTTGTACAGTCCCATTC 1975
Db 64 CTGAGGAACAGTGAATTTCTTCTGCGATTTAAACGCAATTTGTACAGTCCCATTC 5
Qy 1976 CCCC 1979
Db 4 CCCC 1

RESULT 3
AL519829
LOCUS
DEFINITION AL519829 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB005YK05
prime, mRNA sequence.
ACCESSION AL519829
VERSION AL519829.1 GI:12783322
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 920)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
1. .920
```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB005YK05"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 181 a 311 c 282 g 137 t
ORIGIN

Query Match 43.1%; Score 869.8; DB 9; Length 920;
Best Local Similarity 98.0%; Pred. No. 2.9e-163;
Matches 901; Conservative 9; Mismatches 6; Indels 3; Gaps 3;

Qy 873 GGAAGCAGGCGCTCGGGGGCCCCCTG-GACCACCAAGTCCAAGAGGCCCCCAGAGAA 931
Db 1 GGAAGCAGGCGCTCGAGRGCCCCCTGAGACCAACAGGTCCAAAAGAGCCCCAGGAAA 60

Qy 932 CCCGGCAGGCAGGCCCCCGGGCCCTCCCGGTTCAGTCCGGGGGGGGTGGCCCGCT 991
Db 61 CCCGGCAGGCAGGCCCCCGGGCCCTCCCGGTTCAGTCCGGGGGGGGTGGCCCGCT 120

Qy 992 GCGGCTACGTGCTCGCATTTCTTACGCGGGCCCTGCGGCGCCCAACAGGGTTAC 1051
Db 121 GCGGCTACGTGCTCGCATTTCTTACGCGGGCCCTGCGGCGCCCAACAGGGTTAC 180

Qy 1052 GAGTGTCTGCTTCGAGGAGTGTGACCAACGTGGGCAACGCTACGAGGAGCCAGC 1111
Db 181 GAGTGTCTGCTTCGAGGAGTGTGACCAACGCTGGGCAACGCTACGAGGAGCCAGC 240

Qy 1112 GGCAAGTTTACTTGCCTCCATCCAGGCTCTACTTCTTCGCTTACCAGTGTCTATGGC 1171
Db 241 GGCAAGTTTACTTGCCTCCATCCAGGCTCTACTTCTTCGCTTACCAGTGTCTATGGC 300

Qy 1172 GCGGCGAGCGCACAGCATGTGGCGGACCTCATGAAGAACGGACAGTCCGGGCCAGC 1231
Db 301 GCGGCGAGCGCACAGCATGTGGCGGACCTCATGAAGAACGGACAGTCCGGGCCAGC 360

Qy 1232 GCCATTGCTCAGGAGCGGACAGAACTACGCACTACGCAACAGAGCTCATTTGCAAC 1291
Db 361 GCCATTGCTCAGGAGCGGACAGAACTACGCACTACGCAACAGAGCTCATTTGCAAC 420

Qy 1292 CTGAGTGTGGCGAGGAGGTCTTCATCAAGCTGACGCGGGAAGTGCAGCGGGGCAAC 1351
Db 421 CTGAGTGTGGCGAGGAGGTCTTCATCAAGCTGACGCGGGAAGTGCAGCGGGGCAAC 480

Qy 1352 ACCAACAAGTACAGCACCTTCTCGGCTTCATCATCTACCCGACTAGCGGCCCGCC 1411
Db 481 ACCAACAAGTACAGCACCTTCTCGGCTTCATCATCTACCCGACTAGCGGCCCGCC 540

Qy 1412 CCGTGCCCCCGCTCGCCCCCTTCTCTCCCGTCTCACCCACCTCTGCCCCCCCCACCGA 1471
Db 541 CCGTGCCCCCGCTCGCCCCCTTCTCTCCCGTCTCACCCACCTCTGCCCCCCCCACCGA 600

Qy 1472 GCGGCCACCCACCCCTTTGAGAGCTGCGGTGGGGTGGACCTTCCCTTCCCGGAGCG 1531
Db 601 GCGGCCACCCACCCCTTTGAGAGCTGCGGTGGGGTGGACCTTCCCTTCCCGGAGCG 660

Qy 1532 GCCTAAATGGCGGAACCTTCTGCTCAAGGATATAAGTGGCGGAGAGAGAGAGAGACC 1591
Db 661 GCCTAAATGGCGGAACCTTCTGCTCAAGGATATAAGTGGCGGAGAGAGAGAGAGACC 720

Qy 1592 CGGCCAGAGGAGAGAGGAGTTCGGAGGAGTACCCCGCAGCCCAAGTGCCTGGTGGACC 1651
```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB005YK05"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
```

```
Db 721 CGCCAGAGGAGCAGAGC-ACCTCCGGAGGATACCCGACCCCAAGTGCCTGGACC 779
QY 1652 CCATAGGGGACAGAGTCTGCTCTCTTTTGTACAGAGATGGGAGCAGTTTAAATAG 1711
Db 780 CCATAGGGGACAGAGTCTGCTCTCTTTTGTACAGAGATGGGAGCAGTTTAAATAG 839
QY 1712 CGGGACTCAGAGCCAGAAAGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1771
Db 840 CGGGACTCAGAGCCAGAAAGCCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 898
QY 1772 GAAACAGGAGGAGCCCAT 1790
Db 899 GAAACAGGAGGAGGAGCCAT 917

RESULT 4
AL529898 942 bp mRNA linear EST 13-FEB-2001
LOCUS AL529898 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD005YH17 5
DEFINITION prime, mRNA sequence.
ACCESSION AL529898
VERSION AL529898.1 GI:12793391
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr.

FEATURES
source
location/Qualifiers
1..942
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DD005YH17"
/clone_lib="LTI_NFL001_NBC4"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact: Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax: (1) 301 610
8371 Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com"
BASE COUNT 160 a 307 g 112 t 6 others
ORIGIN
Query Match 43.1%; Score 868.6; DB 9; Length 942;
Best Local Similarity 98.8%; Pred. No. 4.9e-163;
Matches 890; Conservative 5; Mismatches 4; Indels 2; Gaps 2;

QY 24 GGACCGCGCCGACAGAGAGCGCGGGCCAGGGTACTAGCAGGAGTGGGGCGCGCGC 83
Db 1 GGACCGCGCCGACAGAGAGCGCGGGCCAGGGTACTAGCAGGAGTGGGGCGCGCGC 60
QY 84 AGGGGTAGCAAGGTGAGTGGTCTTCCCAAGCAGGAGCGCCAAACCTACTAGAGAT 143
Db 61 AGGGGTAGCAAGGTGAGTGGTCTTCCCAAGCAGGAGCGCCAAACCTACTAGAGAT 120
QY 144 CGCGCCGGTGAAGCAGCAGCCAGCTCAGAGCCCGGAGCTCCGAGCGCGGGAGCAG 203
Db 121 CGCGCCGGTGAAGCAGCAGCCAGCTCAGAGCCCGGAGCTCCGAGCGCGGGAGCAG 180
```

```
QY 204 TCCCTCTCCATCAGGGAGTGTCTATCTGCGCAGTCTTGGGACCCAGGACCCGCGGCATC 263
Db 181 TCCCTCTCCATCAGGGAGTGTCTATCTGCGCAGTCTTGGGACCCAGGACCCGCGGCATC 240
QY 264 CTTGAGAGAGCAGCAGTCTGGAGAGCAGGCATCTCAGATCCCTAAGAACACGAGCGTCCG 323
Db 241 CTTGAGAGAGCAGCAGTCTGGAGAGCAGGCATCTCAGATCCCTAAGAACACGAGCGTCCG 300
QY 324 AGAAGCCGCGGATCTCAGGTGCCAGGATCGTTAGGACTGAACGGGAGGGTACTAGAGGA 383
Db 301 AGAAGCCGCGGATCTCAGGTGCCAGGATCGTTAGGACTGAACGGGAGGGTACTAGAGGA 360
QY 384 CCACCTGGTCTTGGACCGTGGAGAGTGCCTTCAGCTAAACCCAGGAGGGCTCCCTCTTG 443
Db 361 CCACCTGGTCTTGGACCGTGGAGAGTGCCTTCAGCTAAACCCAGGAGGGCTCCCTCTTG 420
QY 444 ACGGAGCGCTTGGGAGAGCGGACCCGCGCTTGGAGCCCGCAGAGCAGGGTAAAGGGAG 503
Db 421 ACGGAGCGCTTGGGAGAGCGGACCCGCGCTTGGAGCCCGCAGAGCAGGGTAAAGGGAG 480
QY 504 CGGGGGGACGCGTGGGGAGTGCAGACCCAGGAGCCAAAGGGGGTCAACCGTCTCTGGCC 563
Db 481 CGGGGGGACGCGTGGGGAGTGCAGACCCAGGAGCCAAAGGGGGTCAACCGTCTCTGGCC 540
QY 564 CGCGAGAGCCCGCGCGCGCGCAGCATTTGCGCCCAAGAGTGAAGAGATTTGCTGGCC 623
Db 541 CGCGAGAGCCCGCGCGCGCGCAGCATTTGCGCCCAAGAGTGAAGAGATTTGCTGGCC 600
QY 624 TGGCAGCGTCCGGGCTGAGCGCGCAGAGAGGTTGGCGGCGCGCGCTCGGAGTGGCCA 683
Db 601 TGGCAGCGTCCGGGCTGAGCGCGCAGAGAGGTTGGCGGCGCGCGCTCGGAGTGGCCA 660
QY 684 TGGTCTGCTGCTGTGTGGTGCATCCCGCTGCTGTGTGCACAGTCCCGCGCGCGCAGCG 743
Db 661 TGGTCTGCTGCTGTGTGGTGCATCCCGCTGCTGTGTGCACAGTCCCGCGCGCGCAGCG 720
QY 744 ACTACAGATGCTGGTCTGCTGCGCATGTTGGACCCCGATGCGGATGCGGCGCGCGCTTG 803
Db 721 ACTACAGATGCTGGTCTGCTGCGCATGTTGGACCCCGATGCGGATGCGGCGCGCGCTTG 780
QY 804 -GTCCCGAGCGCGCGCTGCTTCCGTCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCG 862
Db 781 TGTCCCGAGCGCGCGCTGCTTCCGTCGCGCGCGCGCTTCCCGCGCGCGCGCGCGCG 840
QY 863 GGCGCGCGCGGAAAGCAGCGCTTGGGGGGCGCGCGCTTGACACCGAGTCCAGAGGGCGCC 922
Db 841 GGCGCGCGCGGAAAGCAGCGCTTGGGGGGCGCGCGCTTGACACCGAGTCCAGAGGGCGCC 899
QY 923 C 923
Db 900 C 900
```

RESULT 5

AL518440

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL518440 894 bp mRNA linear EST 13-FEB-2001
AL518440 LTI_NFL011_NBC1 Homo sapiens cDNA clone CS0DA009YF06 5
prime, mRNA sequence.

AL518440

AL518440.1 GI:12781933

EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

I (bases 1 to 894)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Db	659	GGGCGCGCGCGTGGAGTGGCCATGTGTGCTGCTGCTGGTGGCCATCCCGTGGTGGT 718
QY	721	GCACAGCTCCGCGGCGCCAGCGACACTACGAGATGCTGGGTGCGTGCCTGCCGATGCTGTCGCA 780
Db	719	GCACAGCTCCGSGSGSSAGCGCACTACGAGANGCTGGTGGCTGCCGCAATGCTGTCGCA 778
QY	781	CCGCGATGGGCCCGCTGGGCCCTG-GTCCGACGGCGCGCTGCTTCGTCGCCCGCCCTTCC 839
Db	779	CCGCGATGGGCCCGTGGCCCTGTGTCGACGGCGCGCTTCTGTCGTCGTCGCCCGCTTCC 838
QY	840	CGCCAGCGGCACAGGAGAGAGTGGCGCGCGCGGAAAGCAGCGCTGCGGGGGGCCCC 896
Db	839	CGCGAGCGCCACAGGAGAGGTGGCGCGCGCGGAAA-SAGGCGCTCGGGGGGCCCC 894
RESULT 6		
AL518439/c	AL518439	LTI_NFL011_NBC1 870 bp mRNA linear EST 13-FEB-2001
LOCUS	AL518439	LTI_NFL011_NBC1 Homo sapiens cDNA clone CS0DA009YF06 3
DEFINITION	prime, mRNA sequence.	
ACCESSION	AL518439	
VERSION	AL518439.1	GI:12781932
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
TITLE	1 (bases 1 to 870)	
JOURNAL	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.	
COMMENT	Full-length cDNA libraries and normalization Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.	
FEATURES	Location/Qualifiers	
source	1..870	
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	/clone="CS0DA009YF06"	
	/clone_lib="LTI_NFL011_NBC1"	
	/sex="male"	
	/tissue_type="neuroblastoma cells"	
	/lab_host="DH10B"	
	/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA	
	was primed with a NotI-oligo(dT) primer. Five prime end	
	enriched, double-stranded cDNA was digested with Not I and	
	cloned into the Not I and Eco RV sites of the pCMVSPORT 6	
	vector. Library is not normalized, but is the control for	
	the normalized libraries. Library was constructed by Life	
	Technologies. Contact : Feng Liang Life Technologies, a	
	division of Invitrogen 9800 Medical Center Drive Rockville,	
	Maryland 20850, USA Fax : (1) 301 610 8371 Email :	
	filiang@lifetech.com URL :	
	http://fulllength.invitrogen.com"	
BASE COUNT	151 a	241 c 301 g 163 t 14 others
ORIGIN		
Query Match	40.9%	Score 824.8; DB 9; Length 870;
Best Local Similarity	95.6%	Pred. No. 2.6e-154;
Matches 832; Conservative	13; Mismatches 25; Indels 0; Gaps 0;	
QY	1113	GCAAGTTTACTTGCCTCCCATGCCAGGGCTGTACTTCTCGCTTACCAGTGCTCATGGCG 1172
Db	870	GCAAGTTTACTTGCCTCCCATGCCAGGGCTCAACTTCTCGCTTAAACAGGCTCATGGCG 811
QY	1173	CGCGCGACGSCACAGCATGTGGCGGCACTCATGAAGACGACAGTCCGGGCCAGCG 1232
Db	810	CGCGCGACGSCACCMGATGTGGCGGCCCTCATGAAGACGACAGTCCGGGCCAGCG 751
QY	1233	CCATTGCTCAGGACGGGCGACAGAATTCAGCTACGCCAGCAACAGCGTCAATTCTGCACC 1292


```
RESULT 8
BE888168
LOCUS      601511679P1 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:3913171 5',
DEFINITION mRNA sequence.
ACCESSION BE888168
VERSION   BE888168.1 GI:10344201
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 387)
           NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
           Email: cgapbs-r@mail.nih.gov
           Tissue Procurement: ATCC.
           cDNA Library Preparation: Life Technologies, Inc.
           DNA Sequencing by: Incyte Genomics, Inc.
           Clone distribution: MGC clone distribution information can be
           found through the I.M.A.G.E. Consortium/LLNL at:
           http://image.llnl.gov
           Plate: L1AM9732 row: n column: 20
           High quality sequence stop: 387.
           Location/Qualifiers
FEATURES             1..387
   source             /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3913171"
                     /clone_lib="NIH_MGC_71"
                     /tissue_type="leiomyosarcoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: uterus; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 2.1 kb."
BASE COUNT          86 a 106 c 149 g 46 t

   Query Match      17.6%; Score 355.8; DB 12; Length 387;
   Best Local Similarity 98.7%; Pred. No. 6.5e-61;
   Matches 380; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

QY 1 AGGAAGTTAGGAGCGGAGAGGACCGCGCCGCA-GAGAGAGCGCGCGCAGGGCT 59
DB 5 AGGAAGTTAGGAGCGGAGAGGACCGCGCCGCAAGAGAGGCGCGCGCGCGAGGCT- 63
QY 60 ACTAGCAGGGACTGGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 119
DB 64 -CTAGCAGGAGTGGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 122
QY 120 AGAGCGCAAACTACTAGGAGATCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 179
DB 123 AGAGCGCAAACTACTAGGAGATCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGG 182
QY 180 GGAGCTCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 239
DB 183 GGAGCTCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 242
QY 240 CTGGGACCCAGGACCGCGGATCCCTGAGAGAGCAGGAGGAGGAGGAGGAGGAGG 299
DB 243 CTGGGACCCAGGACCGCGGATCCCTGAGAGAGCAGGAGGAGGAGGAGGAGGAGG 302
QY 300 GATCCCTAAGAAACAGCCGTCGAGAGGCGCGGATCTCAGGTCGCCAGGATCGTTAGG 359
DB 303 GATCCCTAAGAAACAGCCGTCGAGAGGCGCGGATCTCAGGTCGCCAGGATCGTTAGG 362
QY 360 ACTAAGCGGAGGAGTACTAGAGGAC 384
DB 363 ACTAAGCGGAGGAGTACTAGAGGAC 387
```

```
RESULT 9
BM844470
LOCUS      K-EST0122585 S13KMS5 Homo sapiens cDNA clone S13KMS5-36-C10 5',
DEFINITION mRNA sequence.
ACCESSION BM844470
VERSION   BM844470.1 GI:19200869
KEYWORDS  EST.
SOURCE    human.
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 425)
           Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
           Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
           Kim,Y.S.
           21C Frontier Korean EST Project 2001
           Unpublished (2002)
           Contact: Kim YS
           Genome Research Center
           Korea Research Institute of Bioscience & Biotechnology
           52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
           Tel: +82-42-860-4470
           Fax: +82-42-860-4409
           Email: yongsung@mail.kribb.re.kr
           Plate: 36 row: C column: 10
           High quality sequence stop: 425.
           Location/Qualifiers
FEATURES             1..425
   source             /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="S13KMS5-36-C10"
                     /clone_lib="S13KMS5"
                     /tissue_type="myeloma"
                     /cell_line="KMS-5"
                     /lab_host="Top10F"
                     /note="Vector: pCNS; Site_1: EcoRI; Site_2: NotI; The poly
                     (A)+ RNA was dephosphorylated with bacterial alkaline
                     phosphatase (BAP) and then decapped with tobacco acid
                     pyrophosphatase (TAP). The decapped intact mRNA was
                     ligated with DNA-RNA linker including EcoRI site by
                     treatment of T4 RNA ligase and the first strand cDNA was
                     synthesized from oligo dt-selected mRNA by priming with
                     dt-tailed vector. The dt-tailed vector was adjusted to
                     have about 60nt. The cDNA vector was circularized with E.
                     coli DNA ligase after digestion of EcoRI which site is
                     also included in vector. An RNA strand converted to a DNA
                     strand by Okayama-Berg method. The obtained cDNA vectors
                     were used for transposition of competent cells E. coli
                     Top10F by electroporation method. The cDNA libraries
                     constructed by this method are full-length enriched cDNA
                     library."
BASE COUNT          93 a 119 c 160 g 53 t

   Query Match      17.5%; Score 353.8; DB 14; Length 425;
   Best Local Similarity 96.2%; Pred. No. 1.6e-60;
   Matches 384; Conservative 0; Mismatches 12; Indels 3; Gaps 2;

QY 1 AGGAAGTTAGGAGCGGAGAGGACCGCGCCGCGCAGAGAGAG--GCGCGGCGCCAGGCG 58
DB 27 AGGAAGTTAGGAGCGGAGAGGACCGCGCCGCGCAGAGAGAGAGTGCGCGGCGCAGTGT 86
QY 59 TACTAGCAGGAGCTGGGGCGCGCGGAGGGGTAGCAAGGTAGTGCTGCTTGCCTCCAGAGG 118
DB 87 CTCTAGCAGGAGCTGGGGCGCGCGGAGGGGTAGCAACGTGAGTGTGCTTGCCTCCAGAGG 146
QY 119 CAGAGCGCAAACTACTAGGAGATCGCGCGCGGTGAGCAGCACCAGCTCAGAGGCC 178
DB 147 CAGAGCGCAAACTACTAGGAGATCGCGCGCGGTGAGCAGCACCAGCTCAGAGGCC 206
```


plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an EcoT of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 110 a 209 c 145 g 88 t

ORIGIN

Query Match 15.6%; Score 314.4; DB 13; Length 552;

Best Local Similarity 76.8%; Pred. No. 1.2e-52;

Matches 394; Conservative 0; Mismatches 116; Indels 0; Gaps 0;

QY 1001 GTCCTCGCATGCTTTCTACCGCGCTCGCGCGCCCCACGAGGGTTACGAGTGCTG 1060

Db 8 GTGCCGCGTGGCTTCTACCGCGCTCTCAAGAACCCACGAGGGTTACGAGTACTC 67

QY 1061 CCGTTGCGAGCGTGTGACCAACCTGGCAACGCTACGAGCGACCGCGGCAAGTTT 1120

Db 68 AAGTTTGAGCGTGTGTCACCAACCTAGCAACACTAGCGCGCGCGCAAGTTT 127

QY 1121 ACTTGCCCATGCCAGGGGTCTACTTCTTCGCTTACCAGTCTCATGCGCGCGCGCAC 1180

Db 128 AGTGCACATTTCCCGGCACTTCTTTTCACTTACCATGTCTCATGCGCGCGCGCAC 187

QY 1181 GGCACAGCATGTGGCGGACCTCATGAAGACGACAGGTCCGGGCGCGCCATTGCT 1240

Db 188 GGCACAGTATGTGGCAGACCTCTGCAAGAAATGCCAGGTGCGGCGCGAGTCTATTGCC 247

QY 1241 CAGGACGGGACGACGACTACGCTACGCCACACACGGTCTATTCGACCTGACGCTG 1300

Db 248 CAGGACGGGACGACGACTACGCTACGCCACACACGGTCTATTCGACCTGACGCTG 307

QY 1301 GCGACGAGGTCTTCATCAAGCTGGACGGCGGGAAGTGCACGGCGGCAACACCAAG 1360

Db 308 GCGACGAGGTCTTCATCAAGCTGGATGAGGCAAGACACGCGGCGGCAACACCA 367

QY 1361 TACAGACCTTCTCGGCTTCATCTATCTACCCGCTAGCGCGCGCGCGCGCGCGCG 1420

Db 368 TACAGACCTTCTCGGCTTCATCTATCTACCCGCTAGCTAGCTAGCTAGCTAGCT 427

QY 1421 CGCTGCCCTTCTCTCGGCTCTCAGCCACCTCTCGCGCGCGCGCGCGCGCGCG 1480

Db 428 CCAGTGTCTTACCCCGCGGCTCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 487

QY 1481 CCACCTTTTGAAGCGCTGGC 1500

Db 488 CCACCGCTCGCTGCCCGGC 507

RESULT 13

AQ200834/c

LOCUS

DEFINITION AQ200834 323 bp DNA linear GSS 20-APR-1999

RPCL11-62K10.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-62K10,

DNA sequence.

ACCESSION AQ200834

VERSION AQ200834.1 GI:3613033

KEYWORDS GSS.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,

Berry,K., Granger,D., Suh,E., Wible,C., de Jong,P. and Venter,J.C.

Use of human BAC End Sequences for Sequence-Ready Map Building

Unpublished (1998)

CONTACT: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdadams@tigr.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.tigr.org/tldb/humgen/bac_end_search/bac_end_search.html

Class: BAC ends.

FEATURES

Location/Qualifiers

Source

1..323

/organism="Homo sapiens"

/db_xref="GDB:7523673"

/db_xref="taxon:9606"

/clone="RPCI-11-62K10"

/clone_lib="RPCI-11"

/sex="Male"

/cell_type="Lymphocytes"

/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI;

RPCL11 Human Male BAC Library"

BASE COUNT 61 a 91 c 87 g 84 t

ORIGIN

Query Match 15.5%; Score 313.4; DB 17; Length 323;

Best Local Similarity 98.1%; Pred. No. 1.8e-52;

Matches 317; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1605 AGAGCGACTTCGCGGAGGATCACCGCACCCCAAGTCCGCGTGGACCCCATAGGGCGAGA 1664

Db 323 AGAGCGACTTCAGAGGGATCACCGCACCCCAAGTCCGCGTGGACCCCATAGGGCGAGA 264

QY 1665 GGTCTGGCTTCTCTTTGTACAGATGGGGAGCAGTTTAAATAGCGGGACTCAGAGG 1724

Db 263 GGTAGTGGCTTCTCTTTGTACAGATGGGGAGCAGTTTAAATAGCGGGACTCAGAGG 204

QY 1725 CCCAGAAAGCCGAGGAGGAGCCCGCAGCTTTCGAGGGGAAATACAGAAACAGAGGAG 1784

Db 203 CCCAGAAAGCCGAGGAGGAGCCCGCAGCTTTCGAGGGGAAATACAGAAACAGAGGAG 144

QY 1785 CCCATTTAGGCAAGAGAAGACATTAACAGAGGTAGTGCAGGTTCCTCCTCACACTTTC 1844

Db 143 CCCATTTAGGCAAGAGAAGACATTAACAGAGGTAGTGCAGGTTCCTCCTCACACTTTC 84

QY 1845 TCTCCGCCACCTCTCGTCCCTCTCCACCTTTCAGGCTCAGGCTCCAGCTTTCGACG 1904

Db 83 TCTCCGCCACCTCTCGTCCCTCTCCACCTTTCAGGCTCAGGCTTTCGACGCTTTCGACG 24

QY 1905 CTTCTGTGAAGTGGAGGAACCA 1927

Db 23 CTTCTGTGAAGTGGAGGAACCA 1

RESULT 14

Bi793031

LOCUS

DEFINITION Bi793031 461 bp mRNA linear EST 12-MAR-2002

ie47a04.y1 Melton Normalized Human Islet 4 M4-HIS 1 Homo sapiens

CDNA clone IMAGE:5670030 5' similar to TR:075973 075973 C1Q-RELATED

FACTOR. ; mRNA sequence.

ACCESSION Bi793031

VERSION Bi793031.1 GI:15820756

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

AUTHORS 1 (bases 1 to 461)

Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,

Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,

Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,

Schmitt,A., Theisinger,B., Rutter,E., Ronko,I., Bennett,J., Cardenas

,M., Gibbons,M., McCann,R., Cole,R., Tsagarelshvili,R., Williams,T.

, Jackson,Y. and Bowers,Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Other_ESTs: ie47a04.x1

TITLE

JOURNAL

COMMENT

Db 368 TACAGCACGTTCTCTGGCTTCATCATCTACTCCGACTGAGCTCCCCACGTCTCCCTCCAC 427

```

Matches 376; Conservative 0; Mismatches 114; Indels 0; Gaps 0;
QY 1011 TTGCTTTTACGGGGCTCGCGCGCCCGCCAGAGGGTTACGAGGTGCTGCGCTTCGACG 1070
Db 4 TGGCTTTTACGGGGCTCAAGAACCCCGCCAGAGGGTTACGAGGTACTCAAGTTTGACG 63
QY 1071 AGTGTGTACCAACAGTGGGCAACGCTACGAGCGAGCCAGCGCAAGTTTACTTGGCCCA 1130
Db 64 AGTGTGTACCAACCTAGGCAACAATACGAGCGGCGCCAGCGCAAGTTTACGTGCAACA 123
QY 1131 TGCAGGCGCTCTACTTCTTCGCTTACCAAGTGTCTCATGCGGCGCGGACGCCAGCA 1190
Db 124 TTCCCGGCACTTCTTTCACCTACCATGTCCTCATGCGGCGGCGGACGCCAGTA 183
QY 1191 TGTGGCGGACCTCATGAAGACGGACAGGTCCGGGCCAGCGCCATTGCTCAGGACGCGG 1250
Db 184 TGTGGCGACACCTCTGCAAGATGGCCAGGTGCGGGCCAGTGCTATTGCCAGGACGCGG 243
QY 1251 ACCAGAACTACGACTACGCCAGCAACAGCGTCTATTGCACTGGACGTGGGCGGAGG 1310
Db 244 ACCAGAACTACGACTACGCCAGCAACAGCGTGTATCTGCACTGGACGTGGGCGGAGG 303
QY 1311 TCTTCATCAAGCTGGAGCGGGAAAGTGCACGGGCAACACACACACACACACACCT 1370
Db 304 TCTTCATCAAGCTGGATGGAGGCAAGCACCGGGCAACACACACACACACACACCT 363
QY 1371 TCTCGGCTTCATCATCTACCCGACTGAGCGGGCGCCCGCCCGCTCGCCCGC 1430
Db 364 TCTCGGCTTCATCATCTACCCGACTGAGCTCCCGACGCTCTCCCTCCACCCACGTCCT 423
QY 1431 TTCTCTCCCGTCCGTACCCACCTCTCTGCGCGCGCCACCCGAGGCGCCACCCACCTTTG 1490
Db 424 CACCGCGCGGGTCCCGCTCCGGGCGGCGAGACGATGACTCGCCCGCTCGCCCGCGCTC 483
QY 1491 AGAGCCTGGC 1500
Db 484 GCTGCCCGGC 493

```

Search completed: June 6, 2003, 20:53:01
Job time : 2940 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2003, 16:36:13 ; Search time 72 Seconds
(without alignments)
440.467 Million cell updates/sec

Title: US-10-005-499-378
Perfect score: 1311
Sequence: 1 MVLLLVAILPLVHSSRGA.....VHGNTNKYSTFGFIIPD 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database : A_Geneseq_101002.*

- 1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1311	100.0	238	AAU84377	Novel human secret
2	1286.5	98.1	267	AAU84378	Novel human secret
3	907.5	69.2	255	AAU84379	Murine HSP47 inter
4	903.5	68.9	253	AAU84380	Human polypeptide
5	869.5	66.3	287	AAU84381	Novel human secret
6	662.5	50.5	225	AAU76873	Human CRF-like pro
7	376	28.7	744	AAU39127	Human polypeptide
8	376	28.7	755	AAU40913	Human polypeptide
9	352	26.8	253	AAU38821	Human polypeptide
10	352	26.8	255	AAU40607	Human polypeptide

11	345.5	26.4	638	23	ABJ01025	Human breast speci
12	345	26.3	635	22	AAU78798	Human protein SEQ
13	345	26.3	644	22	AAU79782	Human protein SEQ
14	329.5	25.1	314	23	ABP41487	Human ovarian anti
15	329	25.1	285	21	AAU37344	Murine ACRP30R1M.
16	328	25.0	285	21	AAU30232	A polypeptide desi
17	328	25.0	285	21	AAU54321	Human adipocyte co
18	328	25.0	285	21	AAU50374	Human adipocyte co
19	327	24.9	245	21	AAU43063	Human OREF ORF2827
20	327	24.9	245	22	AAU38852	Human polypeptide
21	327	24.9	277	22	AAU40638	Human polypeptide
22	327	24.9	285	22	AAU76103	Rat skin cell secr
23	327	24.9	285	22	AAU56042	Rat skin cell secr
24	327	24.9	285	23	ABU72242	Rat protein isolat
25	327	24.9	294	21	AAU76039	Rat skin cell prot
26	327	24.9	294	22	AAU55978	Skin cell protein,
27	327	24.9	294	23	ABU72178	Rat protein isolat
28	327	24.9	331	22	AAU19557	Human diagnostic a
29	324.5	24.8	244	21	AAU85144	Human adiponectin
30	322	24.6	244	18	AAU09108	Human complement c
31	322	24.6	245	21	AAU30234	Human Cql c protei
32	322	24.6	245	22	AAU49594	Human Cql c protei
33	322	24.6	245	22	AAU49600	Human Cql c protei
34	320	24.4	244	21	AAU71053	Human APM1 (Adipos
35	319.5	24.4	289	22	AAU50376	Mouse ZACRP7. Mus
36	319.5	24.4	289	23	ABU72306	Adipose most abund
37	318	24.3	244	20	AAU21807	Rat protein isolat
38	318	24.3	244	21	AAU30233	Human adipocyte co
39	318	24.3	244	21	AAU71035	Human APM1 (Adipos
40	318	24.3	244	22	AAU05529	Human OBG3 protein
41	318	24.3	244	22	AAU65828	Human adipocyte co
42	318	24.3	244	22	AAU49598	Human ACRP30 prote
43	318	24.3	244	22	AAU50373	Human adipocyte co
44	318	24.3	244	23	ABU08223	Human apml protein
45	318	24.3	244	23	AAU80254	Human APM1 protein

ALIGNMENTS

RESULT 1
AAU84377
ID AAU84377-standard; Protein; 238 AA.
XX
AC AAU84377;
XX
DT 08-MAY-2002 (first entry)
XX
DE Novel human secreted or membrane-associated protein #16.
XX

Human; secreted protein; membrane-associated protein; hypertension;
inflammatory disorder; neurological disorder; hematopoietic disorder;
skeletal developmental disorder; growth abnormality; autoimmune disorder;
neurodegenerative disorder; nervous system disorder; bacterial infection;
peripheral myelinopathy; viral infection; cancer; obesity; diabetes;
hypotension; sexual development disorder; blood disorder.

OS Homo sapiens.

XX WO200204600-A2.

XX 17-JAN-2002.

XX 12-JUL-2001; 2001WO-US21985.

XX 12-JUL-2000; 2000US-218033P.

XX 21-AUG-2000; 2000US-226517P.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX (GLAX) SMITHKLINE BEECHAM PLC.

XX (GLAX) GLAXO GROUP LTD.

XX Agarwal P, Cogswell JP, Lai Y, Martensen SA, Rizvi SK, Strum JC;

```
PI Smith RF, Xiang Z, Xie Q;
XX WPI; 2002-188468/24.
DR N-PSDB; ABK35597.
XX
XX Novel secreted and membrane-associated polypeptides and polynucleotides
PT encoding the polypeptides, for preventing, treating and ameliorating
PT cancers, mental or sexual developmental disorders, and malignant tumours
PT
XX
XX Claim 1; Page 139-140; 151pp; English.
PS
XX The present invention relates to the isolation of novel human secreted
XX or membrane-associated proteins and the genes encoding them. The
XX sequences of the invention are useful for treating, preventing and
XX ameliorating various diseases such as inflammatory disorders (e.g.
XX asthma), neurological disorders (e.g. dementia), haematopoietic
XX disorders, skeletal developmental disorders, growth abnormalities,
XX neurodegenerative disorders (e.g. Huntington's disease), nervous system
XX disorders, autoimmune disorders (e.g. rheumatoid arthritis),
XX peripheral myelinopathies, viral and bacterial infections,
XX alpha-mannosidosis, diabetes, cancers, malignant tumours, hyper- and
XX hypotension, obesity, bulimia, anorexia, manic depression, delirium,
XX mental retardation, Tourette's syndrome, schizophrenia, growth, mental
XX or sexual development disorders, and dysfunctions of the blood cascade
XX system including those leading to stroke. AAU84362-AAU84389 represent
XX the novel human secreted or membrane-associated proteins of the
XX invention.
XX
XX Sequence 238 AA;
XX
XX Query Match 100.0%; Score 1311; DB 23; Length 238;
XX Best Local Similarity 100.0%; Pred. No. 8.3e-101;
XX Matches 238; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 MVLILLVATPLLVHSSRGPAHYEMLGRCRMVCDPHGPRGPDGAPASVPPPPGAKGEV 60
XX DB 1 MVLILLVATPLLVHSSRGPAHYEMLGRCRMVCDPHGPRGPDGAPASVPPPPGAKGEV 60
XX
XX QY 61 GRRGKAGLRGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 120
XX DB 61 GRRGKAGLRGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 120
XX
XX QY 121 EGYEVLRFDDVTNNGNAYEAASGKFTCPMPGVYFFAYHVLMRGGDGTSMWADLMKNGOV 180
XX DB 121 EGYEVLRFDDVTNNGNAYEAASGKFTCPMPGVYFFAYHVLMRGGDGTSMWADLMKNGOV 180
XX
XX QY 181 RASATAQADQNDYASNSVILHLDVGDEVFIKLDGKVGHTNKYSTFSGFIIPD 238
XX DB 181 RASATAQADQNDYASNSVILHLDVGDEVFIKLDGKVGHTNKYSTFSGFIIPD 238
XX
XX RESULT 2
XX AAU84378
XX ID AAU84378 standard; Protein; 267 AA.
XX
XX AC AAU84378;
XX
XX DT 08-MAY-2002 (first entry)
XX
XX DE Novel human secreted or membrane-associated protein #17.
XX
XX KW Human; secreted protein; membrane-associated protein; hypertension;
XX inflammatory disorder; neurological disorder; haematopoietic disorder;
XX skeletal developmental disorder; growth abnormality; autoimmune disorder;
XX neurodegenerative disorder; nervous system disorder; bacterial infection;
XX peripheral myelinopathy; viral infection; cancer; obesity; diabetes;
XX hypotension; sexual development disorder; blood disorder.
XX
XX OS Homo sapiens.
XX
XX PN WO200204600-A2.
XX
XX AC
```

```
PD
XX
XX 17-JAN-2002.
XX
XX 12-JUL-2001; 2001WO-US21985.
XX
XX 12-JUL-2000; 2000US-218033P.
XX
XX 21-AUG-2000; 2000US-226517P.
XX
XX (SMIK ) SMITHKLINE BEECHAM CORP.
XX PA (SMIK ) SMITHKLINE BEECHAM PLC.
XX PA (GLAX ) GLAXO GROUP LTD.
XX
XX Agarwal P, Cogswell JP, Lai Y, Martensen SA, Rizvi SK, Strum JC;
XX Smith RF, Xiang Z, Xie Q;
XX
XX WPI; 2002-188468/24.
XX N-PSDB; ABK35598.
XX
XX Novel secreted and membrane-associated polypeptides and polynucleotides
XX encoding the polypeptides, for preventing, treating and ameliorating
XX cancers, mental or sexual developmental disorders, and malignant tumours
XX
XX
XX Claim 1; Page 140; 151pp; English.
XX
XX The present invention relates to the isolation of novel human secreted
XX or membrane-associated proteins and the genes encoding them. The
XX sequences of the invention are useful for treating, preventing and
XX ameliorating various diseases such as inflammatory disorders (e.g.
XX asthma), neurological disorders (e.g. dementia), haematopoietic
XX disorders, skeletal developmental disorders, growth abnormalities,
XX neurodegenerative disorders (e.g. Huntington's disease), nervous system
XX disorders, autoimmune disorders (e.g. rheumatoid arthritis),
XX peripheral myelinopathies, viral and bacterial infections,
XX alpha-mannosidosis, diabetes, cancers, malignant tumours, hyper- and
XX hypotension, obesity, bulimia, anorexia, manic depression, delirium,
XX mental retardation, Tourette's syndrome, schizophrenia, growth, mental
XX or sexual development disorders, and dysfunctions of the blood cascade
XX system including those leading to stroke. AAU84362-AAU84389 represent
XX the novel human secreted or membrane-associated proteins of the
XX invention.
XX
XX Sequence 267 AA;
XX
XX Query Match 98.1%; Score 1286.5; DB 23; Length 267;
XX Best Local Similarity 89.1%; Pred. No. 1e-98;
XX Matches 238; Conservative 0; Mismatches 0; Indels 29; Gaps 1;
XX
XX QY 1 MVLILLVATPLLVHSSRGPAHYEMLGRCRMVCDPHGPRGPDGAPASVPPPPGAKGEV 60
XX DB 1 MVLILLVATPLLVHSSRGPAHYEMLGRCRMVCDPHGPRGPDGAPASVPPPPGAKGEV 60
XX
XX QY 61 GRRGKA-----GLRGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 91
XX DB 61 GRRGKAGLRGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 120
XX
XX QY 92 PFGPGGVPAAAGYVPRIAFYAGLRPHEGVLFDDVTNNGNAYEAASGKFTCPMP 151
XX DB 121 PFGPGGVPAAAGYVPRIAFYAGLRPHEGVLFDDVTNNGNAYEAASGKFTCPMP 180
XX
XX QY 152 GYFFAYHVLMRGGDGTSMWADLMKNGQVRAIAQADQNDYASNSVILHLDVGDEVF 211
XX DB 181 GYFFAYHVLMRGGDGTSMWADLMKNGQVRAIAQADQNDYASNSVILHLDVGDEVF 240
XX
XX QY 212 IKLDGKVGHTNKYSTFSGFIIPD 238
XX DB 241 IKLDGKVGHTNKYSTFSGFIIPD 267
XX
XX RESULT 3
XX AAG64212.
XX ID AAG64212 standard; Protein; 255 AA.
XX
XX AC AAG64212;
```

AL: AAG64212

XX 19-SEP-2001 (first entry)
 XX Murine HSP47 interacting protein, #2.
 DE Murine HSP47 interacting protein, #2.
 KW Murine; heat shock protein interacting protein; HSP47.
 XX Mus sp.
 OS JP2001145493-A.
 PN 29-MAY-2001.
 PD 19-NOV-1999; 99JP-0330631.
 XX 19-NOV-1999; 99JP-0330631.
 XX (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
 XX WPI; 2001-395263/42.
 DR N-PSDB; AAH48068.
 XX Using the two-hybrid screening method to prepare proteins which
 PT interact with the heat shock protein HSP47.
 XX Claim 7; Page 23; 26pp; Japanese.
 XX The present invention relates to a method for preparing a protein which
 CC interacts with the heat shock protein HSP47. The method involves the
 CC two-hybrid screening method using the HSP47 gene and a mammalian cDNA
 CC library. The present sequence is the protein sequence for a murine HSP47
 CC interacting protein, which was used in the present method. The HSP47
 CC interacting proteins are useful for the diagnosis and treatment of
 CC diseases caused by an increase or decrease in activity of HSP47.
 XX Sequence 255 AA;

Query Match 69.2%; Score 907.5; DB 22; Length 255;
 Best Local Similarity 70.2%; Pred. No. 2.2e-67;
 Matches 179; Conservative 18; Mismatches 41; Indels 17; Gaps 7;
 QY 1 MVLILLVAIPLLVHSSRGPAHYEMGLRCRMVCDPH-GPRGPG-PPGPPGPPGPPGPP 54
 DB 1 MVLILLVILPVLVSSAGTSAHYEMGLTCRMVCDPYGCTKAPSTAATPDRLQSLPTFIQ 60
 QY 55 GAKGEVGRGKAGLRGP-PPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 105
 DB 61 GPKGEAGRGKAGRGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 120
 QY 106 Y--VPRTAFYAGLRPHGEYEVLRFDVVTVNNGAYEASGKFTCPMPGVYFFAYHVLNR 163
 DB 121 YSTVPKIAFYAGLRQHEGYEVLFKFDVVTVNNGHNDPTTKFTCSIPGIYFFTVHVLNR 180
 QY 164 GGDGTSWADLMKNGQVRASIAQADADQNYDASNVILHLDVGEVFIKLDGKGVHGN 223
 DB 181 GGDGTSWADLMKNGQVRASIAQADADQNYDASNVILHLDVGEVFIKLDGKGVHGN 240
 QY 224 TNKYSTSGFIYIPD 238
 DB 241 NKKYSTSGFIYAD 255

RESULT 4
 ABB53290
 ID ABB53290 standard; Protein; 255 AA.
 XX ABB53290;
 AC ABB53290;
 XX 12-FEB-2002 (first entry)
 XX Human polypeptide #30.
 XX Human; nootropic; neuroprotective; anticonvulsant; antidepressant;

KW neuroleptic; tranquiliser; antiarrhythmic; cardiant; antiasthmatic;
 KW antiinflammatory; antilipaeamic; hepatotropic; virucide; antidiabetic;
 KW nephrotropic; anorectic; cytostatic; vaccine; neurological disease;
 KW cardiovascular disease; respiratory disease; liver disease;
 KW renal disease; skeletal muscle disease; gastrointestinal disease;
 KW placental disease; testicular cancer; male fertility; pancreatic disease.
 XX Homo sapiens.
 OS WO200181363-A1.
 PN 01-NOV-2001.
 PD 26-APR-2001; 2001WO-US13360.
 XX 27-APR-2000; 2000US-199963P.
 PR 11-MAY-2000; 2000US-20336P.
 PR 25-MAY-2000; 2000US-207087P.
 PR 26-MAY-2000; 2000US-207546P.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX Agarwal P, Murdock PR, Rizvi SK, Smith RF, Xiang Z, Kabnick KS;
 PI Lai Y, Xie Q;
 DR WPI; 2002-041392/05.
 DR N-PSDB; ABA90355.
 XX Novel polypeptides and polynucleotides useful as a vaccine for
 PT preventing and treating diseases associated with the polypeptide, e.g.
 PT Alzheimer's disease, dyslipidemia, obesity, diabetes, infertility,
 PT asthma, amnesia
 XX Claim 1; Page 101; 116pp; English.
 XX The invention relates to an isolated polypeptide comprising a 277, 480,
 CC 583, 581, 628, 424, 638, 229, 310, 841, 241, 369, 382, 185, 586, 1026,
 CC 844, 782, 262, 394, 471, 485, 286, 533, 495, 350, 619, 490, 462, 255,
 CC 784, 252, 593, 472, 607, 781, 640, 686 or 154 amino acid sequence as
 CC given in the specification. The polypeptides, modulators of the
 CC polypeptides and antibodies against the polypeptides are useful for
 CC treating diseases such as neurological and psychiatric diseases
 CC including Alzheimer's, paraneoplastic, Huntington's disease,
 CC myotonic dystrophy, anorexia and depression; cardiovascular diseases
 CC including congestive heart failure, Hodgkin's disease and myocardial
 CC infarction; respiratory diseases including asthma, chronic obstructive
 CC pulmonary disease, cystic fibrosis and adult respiratory distress
 CC syndrome; liver diseases including hypercholesterolaemia, cirrhosis,
 CC viral and nonviral hepatitis, Type II diabetes mellitus, and impaired
 CC glucose tolerance; renal disease including renal failure, acute tubular
 CC necrosis and glomerulonephritis; skeletal muscle diseases including
 CC Eulenburg's disease, hypoglycaemia and obesity; gastrointestinal
 CC diseases including myotonia congenita and intestinal obstruction; lymph
 CC diseases including lymphangiectasia; diseases of placenta including
 CC choriocarcinoma; diseases of testes including testicular cancer,
 CC male reproductive diseases including low testosterone and male
 CC infertility; and diseases of pancreas including diabetic ketoacidosis,
 CC Type 1 and 2 diabetes and obesity. The present sequence is a
 CC polypeptide of the invention.
 XX Sequence 255 AA;

Query Match 68.9%; Score 903.5; DB 23; Length 255;
 Best Local Similarity 69.8%; Pred. No. 4.8e-67;
 Matches 178; Conservative 18; Mismatches 42; Indels 17; Gaps 7;
 QY 1 MVLILLVAIPLLVHSSRGPAHYEMGLRCRMVCDPH-GPRGPG-PPGPPGPPGPP 54
 DB 1 MVLILLVILPVLVSSAGTSAHYEMGLTCRMVCDPYGCTKAPSTAATPDRLQSLPTFIQ 60
 QY 55 GAKGEVGRGKAGLRGP-PPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 105

The invention relates to an isolated novel secreted polypeptide (I) and polynucleotide (II). (I) and (II) are useful for treating cancer, autoimmune diseases, wound healing disorder, infections, hematopoietic disorders, inflammatory disorders, infertility, neurological and psychiatric diseases, cardiovascular diseases, respiratory diseases, renal diseases, or gastrointestinal diseases. These may also be used to treat diseases, abnormalities and disorders caused by abnormal expression, production, function and/or metabolism of the genes, as vaccines for inducing immunological response in a mammal, and in screening methods for detecting the effect of added compounds on the production of mRNA and polypeptide in cells. The polypeptides can be used as immunogens to produce antibodies immunospecific for the polypeptides, and to identify membrane-bound or soluble receptors. The polynucleotides may be used as diagnostic reagents, in chromosome localisation studies, and in tissue expression studies. The present sequence represents the amino acid sequence of novel human secreted protein #6.

CC and infection related to a human T-cell leukemia virus (HTLV) type I
CC protein, LP231, which has sequence homology to human cerebellin, and its
CC associated nucleic acid. The polypeptides and polynucleotides are useful
CC for treating or preventing neurological disorders associated with the
CC inappropriate expression of cerebellin-2 proteins and disruption of the

The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAW42213) with neurotrophic, immunosuppressant and cytostatic activity. The polynucleotides are used in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: Immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening assays for receptor activity, arthritis and inflammation, leukaemias C.N.S disorders.

Note: The sequence data for this patent did not form part of the prior specification.

Query Match 28.7%: Score 376: DB 22: Length 744:

XX

DR N-PSDB; AA158283.

Db 64 LPGLAGDHGEFGEKGDPCIPGNPKVGPCKGPMGPKGGPGAPGAPGPKGESGDYKA---TQ 120

Qy 109 RIAFYA-----GLRRPHEGYEVLRFDDVTVNNGNAYEAASGKFTCPMPGVYFFAYHVL 162

Db 121 KIAFSATRTINVPLRRD---QTIRFDHVITNMNNYEPKSGKFTCKVPGLYTYTHASS 176

Qy 163 RGGDGTSMWADLMKNGQVRASAIADADQNDY---YASNSVILHLDVGDVEFIKLDGGKV 219

Db 177 RG----NLCVNLMR-GRERAKVVTFCDYAYNTFQVTTGGMVLEQENYFLOATDKNS 231

Qy 220 HGGNTNKYSTFSGFIIYPD 238

Db 232 LLGMEGANSIFSGLFLFPD 250

RESULT 10

AAM40607

ID AAM40607 standard; Protein; 255 AA.

XX AC AAM40607;

XX DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 5538.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US34263.

XX PR 21-JAN-2000; 2000US-0498725.

XX PR 25-APR-2000; 2000US-052317.

XX PR 09-JUL-2000; 2000US-0598042.

XX PR 19-JUL-2000; 2000US-0620312.

XX PR 03-AUG-2000; 2000US-0653450.

XX PR 14-SEP-2000; 2000US-0662191.

XX PR 19-OCT-2000; 2000US-0693036.

XX PR 29-NOV-2000; 2000US-0727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;

XX PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;

XX PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;

XX DR WPI; 2001-442253/47.

XX DR N-PSDB; AAI59763.

XX PT Novel nucleic acids and polypeptides, useful for treating disorders

XX PT such as central nervous system injuries -

XX XX

XX Example 2; SEQ ID NO 5538; 10078pp; English.

XX CC The invention relates to human nucleic acids (AAI57798-AAI61369) and

XX CC the encoded polypeptides (AAM38642-AAM42213) with nootropic,

XX CC immunosuppressant and cytostatic activity. The polynucleotides are useful

XX CC in gene therapy. A composition containing a polypeptide or polynucleotide

XX CC of the invention may be used to treat diseases of the peripheral nervous

XX CC system, such as peripheral nervous injuries, peripheral neuropathy and

XX CC localised neuropathies and central nervous system diseases, such as

XX CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

XX CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

XX CC utilisation of the activities such as: Immune system suppression,

CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders.

CC Note: The sequence data for this patent did not form part of the printed

CC specification.

XX SQ Sequence 255 AA;

Query Match 26.8%; Score 352; DB 22; Length 255;

Best Local Similarity 35.9%; Pred No. 2.1e-21;

Matches 93; Conservative 34; Mismatches 90; Indels 42; Gaps 10;

Qy 2 VLLLLVAIPLLVHSSRGPAHYEMLRCRMVCDPHGPRGPGDPGAPASVPPPPGAK---- 57

Db 14 VLMLLLGLLIDISQ---AQLSCTGPPAIPGIPGTPGDPGQPGT-----PGIKGEKG 65

Qy 58 -----GEVGRGKAGLRGPPGPPGPRGPGEPGRGPPGPPGPPGAPVPPPPGAK---- 108

Db 66 LPGLAGDHGEFGEKGDPCIPGNPKVGPCKGPMGPKGGPGAPGAPGPKGESGDYKA---TQ 122

Qy 109 RIAFYA-----GLRRPHEGYEVLRFDDVTVNNGNAYEAASGKFTCPMPGVYFFAYHVL 162

Db 123 KIAFSATRTINVPLRRD---QTIRFDHVITNMNNYEPKSGKFTCKVPGLYTYTHASS 178

Qy 163 RGGDGTSMWADLMKNGQVRASAIADADQNDY---YASNSVILHLDVGDVEFIKLDGGKV 219

Db 179 RG----NLCVNLMR-GRERAKVVTFCDYAYNTFQVTTGGMVLEQENYFLOATDKNS 233

Qy 220 HGGNTNKYSTFSGFIIYPD 238

Db 234 LLGMEGANSIFSGLFLFPD 252

RESULT 11

ABJ01025

ID ABJ01025 standard; Protein; 638 AA.

XX AC ABJ01025;

XX DT 05-SEP-2002 (first entry)

XX DE Human breast specific protein SEQ ID NO: 108.

XX KW Human; breast specific gene; breast cancer; gene therapy; breast disease; cytostatic.

XX OS Homo sapiens.

XX PN WO200240672-A2.

XX PD 23-MAY-2002.

XX PF 20-NOV-2001; 2001WO-US45079.

XX PR 20-NOV-2000; 2000US-249998P.

XX PR 22-NOV-2000; 2000US-252563P.

XX XX

XX PA (DIAD-) DIADEXUS INC.

XX PI Salceda S, Macina RA, Recipon H, Cafferkey R, Sun.Y, Liu C;

XX PI WPI; 2002-500220/53.

XX DR

XX PT Novel breast-specific polypeptides and polynucleotides encoding

XX PT polypeptide, useful for identifying, diagnosing, monitoring, staging,

XX PT imaging and treating breast cancer and non-cancerous disease states in

XX PT breast -

XX PS Claim 11; Page 225-227; 243pp; English.

XX CC The present invention provides human breast specific coding sequences and

XX CC proteins. These are useful for detecting breast tissue and for detecting

KW	PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection; inflammatory condition; immune disorder; blood disorder;
KW	cardiovascular disorder; respiratory disorder; neurological disorder;
KW	gastrointestinal disorder; urinary system disorder; drug screening;
KW	gene therapy; chromosome mapping; forensic analysis;
KW	antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW	antiinflammatory; gynaecological; reproductive.
XX	
OS	Homo sapiens.
XX	
PN	WO200200677-A1.
XX	
XX	03-JAN-2002.
XX	
XX	07-JUN-2001; 2001WO-US18569.
XX	
XX	07-JUN-2000; 2000US-209467P.
PR	
XX	(HUMA-) HUMAN GENOME SCI INC.
PA	
XX	Birse CE, Rosen CA;
PI	
XX	WPI; 2002-147878/19.
DR	
DR	N-PSDB; ABQ54564.
XX	
XX	Isolated nucleic acid molecules encoding novel ovarian polypeptides, useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian cancer), immune disorders, cardiovascular disorders and neurological diseases -
PT	
PT	
XX	
XX	Claim 11; SEQ ID NO 2619; 2922pp; English.
XX	
CC	The invention relates to 2175 novel human ovarian antigens (ABP41054-ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also encompasses polypeptides 90% identical and polynucleotides 95% identical to the sequences of the invention. The invention additionally relates to recombinant vectors and host cells comprising human ovarian antigen polynucleotides, antibodies against human ovarian antigens, and the use of ovarian antigen polynucleotides and polypeptides in diagnosing, treating, prognosing or preventing various ovary and/or breast-related disorders. Such conditions include ovarian cancer and breast cancer, and metastatic tumours of ovarian or breast origin, reproductive system disorders (e.g., infertility, disorders of pregnancy, anovulation, polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and vaginitis), immune disorders (e.g., congenital and acquired immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus), blood-related disorders (e.g., anaemia), cardiovascular disorders, respiratory disorders, neurological disorders, gastrointestinal disorders and urinary system disorders. Ovarian antigen polypeptides and polynucleotides may also be used in screening for compounds which modulate ovarian antigen expression or activity. The polynucleotides may further be used for gene therapy, chromosome mapping, in the identification of individuals and in forensic analysis, and the polypeptides may be used as food additives or to prepare antibodies useful in disease diagnosis, drug targeting and phenotyping. The present sequence represents a human ovarian antigen of the invention.
CC	Note: The sequence data for this patent did not form part of the printed specification, and was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
CC	
XX	
XX	Sequence 314 AA;
XX	
XX	Query Match 25.1%; Score 329.5; DB 23; Length 314;
XX	Best Local Similarity 33.6%; Pred. No. 2e-19;
XX	Matches 86; Conservative 37; Mismatches 83; Indels 50; Gaps 10;
XX	
QY	26 GRCWRCDPHGRG-PGPDGAPA-----SVPPPPGA 56
DB	59 GSPQLVCSLEPGQPPGAPGAPSGMWRMGFPFGDKDQDCHDGRDSCSEEGPPRTGN 118
QY	57 KGEVGRRGKAGLRGPPGPPGRGPPGPPGPPGPG-----PGGVAPAAAYVP 108

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2003, 16:36:54 ; Search time 21 Seconds
(without alignments)
470.065 Million cell updates/sec

Title: US-10-005-499-378

Perfect score: 1311

Sequence: 1 MVLVLLVAIPLLVHSSRGPA.....VHGGNTNKYSTFGFIYPD 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	918	70.0	258	1	C1RF_MOUSE
2	917	69.9	258	1	C1RF_HUMAN
3	907.5	69.2	255	1	GLIC_MOUSE
4	397	30.3	743	1	CA18_MOUSE
5	388.5	29.6	744	1	CA18_RABIT
6	376	28.7	680	1	CA1A_MOUSE
7	376	28.7	744	1	CA18_HUMAN
8	362.5	27.7	680	1	CA1A_HUMAN
9	360.5	27.5	674	1	CA1A_CHICK
10	360	27.5	674	1	CA1A_BOVIN
11	352	26.8	251	1	C1QB_HUMAN
12	345	26.3	635	1	CA28_HUMAN
13	328	25.0	285	1	COT2_HUMAN
14	327	24.9	245	1	C1QC_HUMAN
15	318	24.3	244	1	APML_HUMAN
16	317.5	24.2	247	1	APML_MOUSE
17	316.5	24.1	289	1	COT7_HUMAN
18	316	24.1	246	1	C1QC_MOUSE
19	316	24.1	253	1	C1QB_RAT
20	314	24.0	245	1	C1QA_HUMAN
21	302.5	23.1	253	1	C1QB_MOUSE
22	302	23.0	245	1	C1QA_MOUSE
23	300	22.9	243	1	COT5_HUMAN
24	287	21.9	246	1	COT3_HUMAN
25	285.5	21.8	419	1	COLE_LEPMA
26	268.5	20.5	215	1	HP25_TAMSI
27	261.5	19.9	215	1	HP27_TAMSI
28	243	18.5	224	1	CERL_RAT
29	239	18.2	278	1	COT6_HUMAN
30	231	17.6	193	1	CERB_MOUSE
31	230.5	17.6	201	1	CERL_HUMAN
32	229	17.5	193	1	CERB_HUMAN
33	222.5	17.0	196	1	HP20_TAMSI

34	216.5	16.5	281	1	COT1_HUMAN
35	214.5	16.4	1453	1	CA11_CHICK
36	211.5	16.1	1464	1	CA11_HUMAN
37	210.5	16.1	1460	1	CA11_CANFA
38	209.5	16.0	675	1	CA39_CHICK
39	208	15.9	453	1	CA44_BOVIN
40	208	15.9	1466	1	CA13_HUMAN
41	207	15.8	1527	1	CA1H_MOUSE
42	206.5	15.8	1459	1	CA12_MOUSE
43	206.5	15.8	1516	1	CA1H_HUMAN
44	206	15.7	329	1	COT4_HUMAN
45	205.5	15.7	779	1	CA11_BOVIN

ALIGNMENTS

RESULT 1
C1RF_MOUSE STANDARD; PRT; 258 AA.
AC O88992;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Clq-related factor precursor.
GN C1QRF OR CRF.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99097006; PubMed=9878755;
RA Berube N.G., Swanson X.H., Bertram M.J., Kittle J.D., Didenko V.,
RA Baskin D.S., Smith J.R., Pereira-Smith O.M.;
RT "Cloning and characterization of CRF, a novel Clq-related factor,
RT expressed in areas of the brain involved in motor function.";
RL Brain Res. Mol. Brain Res. 63:233-240(1999).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN BRAINSTEM. MORE ABUNDANT IN AREAS
CC OF THE NERVOUS SYSTEM INVOLVED IN MOTOR FUNCTION, SUCH AS THE
CC PURKINJE CELLS OF THE CEREBELLUM, THE ACCESSORY OLIVARY NUCLEUS,
CC THE PONS AND THE RED NUCLEUS.
CC -!- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF095155; AAC64187.1; -.
CC MGD; MGI:1344400; Clqrf.
CC InterPro; IPR001073; Clq.
CC InterPro; IPR000087; Collagen.
CC Pfam; PF00386; Clq; 1.
CC Pfam; PF01391; Collagen; 1.
CC PRINTS; PR00007; COMPLEMENTC1Q.
CC SMART; SM00110; C1Q; 1.
CC PROSITE; PS01113; C1Q; 1.
CC Collagen; Signal.
CC SIGNAL 1 16 POTENTIAL.
CC CHAIN 17 258 C1Q-RELATED FACTOR.
CC DOMAIN 67 115 COLLAGEN-LIKE.
CC DOMAIN 123 258 C1Q.
CC SEQUENCE 258 AA; 26485 MW; F776E2D206EBF763 CRC64;
Query Match 70.0%; Score 918; DB 1; Length 258;
Best Local Similarity 68.6%; Pred. No. 4.8e-53;
Matches 177; Conservative 16; Mismatches 45; Indels 20; Gaps 4;

Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohtsuki S.,
Hayashizaki Y.;
"Functional annotation of a full-length mouse cdna collection.";
Nature 409:685-690(2001).
-!- FUNCTION: MAJOR COMPONENT OF THE DESCEMET'S MEMBRANE (BASEMENT
MEMBRANE) OF CORNEAL ENDOTHELIAL CELLS.
-!- SUBUNIT: MAY FORM HOMOTRIMERS, OR HETEROTRIMERS IN ASSOCIATION
WITH ALPHA 2(VIII) TYPE COLLAGENS.
-!- TISSUE SPECIFICITY: HIGH LEVELS IN CALVARIUM, EYE & SKIN OF
NEWBORN MICE; ALSO IN VARIOUS EPITHELIAL, ENDOTHELIAL AND
MESENCHYMAL CELLS.
-!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPLETIDE REPEATING
UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
-!- SIMILARITY: STRONG, TO ALPHA 2 TYPES VIII AND X COLLAGENS.
-!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
EMBL; X66976; CAA47387.1; -
EMBL; X66977; CAA47387.1; JOINED.
EMBL; AK018742; BAB31383.1; -
PTR: S23779; S23779.

DR INLEIFLO; IFR001073; Clq.
DR InterPro: IPR000087: Collagen

[illegible]

CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X67348; CAA47763.1; -
CC EMBL: X65121; CAA46237.1; -
CC EMBL: X63013; CAA44741.1; -
CC EMBL: X21610; CAA79736.1; -
CC PIR: S28807; S28807.
CC PIR: S31216; S31216.
CC PIR: S22215; S22215.
CC MGD: MGI:88445; Col10a1.
CC InterPro: IPR001073; Clq.
CC InterPro: IPR000087; Collagen.
CC Pfam: PF00386; Clq; 1.
CC Pfam: PF01391; Collagen; 9.
CC PRINTS: PR000007; COMPLENMTCIQ.
CC ProDom: PD000007; Collagen; 2.
CC SMART: SM00110; C1Q; 1.
CC PROSITE: PS01113; C1Q; 1.
CC Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
CC Cartilage; Collagen; Signal.
CC SIGNAL 1 18
CC CHAIN 19 680
CC DOMAIN 19 56
CC DOMAIN 57 519
CC DOMAIN 520 680
CC DOMAIN 545 680
CC CONFLICT 248 248
CC CONFLICT 286 286
CC CONFLICT 306 306
CC CONFLICT 417 417
CC CONFLICT 451 451
CC CONFLICT 451 451
CC CONFLICT 500 500
CC CONFLICT 567 567
CC CONFLICT 569 569
CC CONFLICT 571 571
CC CONFLICT 635 635
CC SEQUENCE 680 AA; 56775 MW; FE984CA99AF708E2 CRC64;
CC -----
CC Query Match 28.7%; Score 376; DB 1; Length 680;
CC Best Local Similarity 37.4%; Pred. No. 1.3e-17;
CC Matches 91; Conservative 31; Mismatches 77; Indels 44; Gaps 9;
CC 36 GPRG---PGPG--APASVPP-----PGAKGEVG---RRGKAGLRGPPGPPRG 79
CC 440 GRGEPGIPGTGTPGPGPGKGDGPNPGAPGAPGATKGLNGTGPDPGPPRG 499
CC 80 PPGEPGRGPPGPGPGGVAP---AAGYVPR-----AFVAGLR 117
CC 500 HSGEPGLGPPGPPGPGQGVMPDGFIRKAGPRUSGMPLVSNHGVTGMPVSFTVILS 559
CC 118 RPHEGYEV-LRFDDVVTVNGVYAEASGFTCPMPGVYFAYHVLMRGSDGTSMWADLMK 176
CC 560 KAYPAVGAPIPDEILYNNQHYDPSRGIFTKIPGIYFYSVHVHK---GTHVWGLYK 616
CC 177 NQGVRSATAQADQNDYASNSVILHLDVGDEVFTKLDGKVGHGNTKNY---STFSGPI 234
CC 617 NGTPTMYTDEYSKGYLDQASGSAIMELTENDQVLMQLPNAESNGLYSSEYVHSSFGSL 676
CC 235 : 1
CC 677 VAP 679

RESULT 7
CA18_HUMAN STANDARD; PRT; 744 AA.
ID CA18_HUMAN Q96D07;
AC P27558; Q96D07;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Collagen alpha 1(VIII) chain precursor (Endothelial collagen).
GN COL8A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91231001; PubMed=2029894;
RA Muragaki Y., Mattei M.-G., Yamaguchi N., Olsen B.R., Ninomiya Y.;
RT "The complete primary structure of the human alpha 1 (VIII) chain and
RT assignment of its gene (COL8A1) to chromosome 3";
RL Eur. J. Biochem. 197:615-622(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR COMPONENT OF THE DESCMET'S MEMBRANE (BASEMENT
CC MEMBRANE) OF CORNEAL ENDOTHELIAL CELLS.
CC -1- SUBUNIT: MAY FORM HOMOTRIMERS, OR HETEROTRIMERS IN ASSOCIATION
CC WITH ALPHA 2(VIII) TYPE COLLAGENS.
CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- MISCELLANEOUS: FOUR CONSECUTIVE GLY-PRO-PRO TRIPLETS ARE PRESENT
CC AT THE C-TERMINUS OF THE TRIPLE-HELICAL REGION. THESE MAY PROVIDE
CC THE HIGH THERMAL STABILITY OF THIS REGION.
CC -1- SIMILARITY: STRONG, TO ALPHA 2 TYPES VIII AND X COLLAGENS.
CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X57527; CAA40748.1; -
CC EMBL: BC013581; AAH13581.1; -
CC PIR: S15435; S15435.
CC Genew: HGNC:2215; COL8A1.
CC MIM: 120251.
CC InterPro: IPR001073; Clq.
CC InterPro: IPR000087; Collagen.
CC Pfam: PF00386; Clq; 1.
CC Pfam: PF01391; Collagen; 8.
CC PRINTS: PR000007; COMPLENMTCIQ.
CC ProDom: PD000007; Collagen; 1.
CC SMART: SM00110; C1Q; 1.
CC PROSITE: PS01113; C1Q; 1.
CC Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
CC Glycoprotein; Cell adhesion; Collagen; Signal.
CC SIGNAL 1 28
CC CHAIN 29 744
CC DOMAIN 29 117
CC DOMAIN 118 571
CC DOMAIN 572 744
CC DOMAIN 609 744
CC CONFLICT 262 262
CC CONFLICT 297 297
CC CONFLICT 344 344
CC CONFLICT 362 362
CC CONFLICT 388 388
CC CONFLICT 454 454
CC COLLAGEN ALPHA 1(VIII) CHAIN.
CC NONHELICAL REGION (NC2).
CC TRIPLE-HELICAL REGION (COL1).
CC NONHELICAL REGION (NC1).
CC C1Q.
CC P -> L (IN REF. 1).
CC P -> R (IN REF. 1).
CC P -> A (IN REF. 1).
CC A -> S (IN REF. 1).
CC P -> S (IN REF. 1).
CC L -> F (IN REF. 1).

```
FT CONFLICT 464 464 A -> H (IN REF. 1).
FT CONFLICT 601 601 Y -> T (IN REF. 1).
FT CONFLICT 631 631 A -> G (IN REF. 1).
SQ SEQUENCE 744 AA; 73364 MW; 2BC1B0955DE2C9A3 CRC64;

Query Match
Best Local Similarity 37.2%; Pred. No. 1.4e-17; Length 744;
Matches 90; Conservative 31; Mismatches 77; Indels 44; Gaps 10;

QY 34 PHGPRG-PGPDGAPASVPPFPKAGKEGVRGKAGLRGPPGPPGPGGRCRPGPGP 92
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 508 PTGPPGIPGKPEGL--PGPGFPG-IGKPGVAGLHPPGKPGALGPGQPGPLPGPGP 564
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 93 PG-PCGGV-----APAAGYVPRIAF-YAGLRPH-----EGYEV----- 125
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 PGPPGPPAVMPTPPPGQGVLPDMGLGIDGVPKPPHAYGAKKNGKNGSPAYEMPFAETA 624
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 126 -----LRFDVTVNGNAYEAASGKTCMPGVFFAYHVLVRGGDGTSMWADLMKN 177
   ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 625 PFPVPGAPVKFNKLLYNGRQNPQTGFTCEVPGVYFAYHVHCKGG---NWVALFRN 681
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 178 GOVRASAIQAADQNDYASNVILHLVDGDEVIKLDGKGVHGGNTKV--STFSGFII 235
   ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::| ::|
Db 682 NEPVWYTYDEYKKGFLDQAGSNAVLLLRPGDRVFLQMPSEQAAGLYAGQIVHSSFGYLL 741
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 236 YP 237
   ||
Db 742 YP 743

RESULT 8
CAJA_HUMAN STANDARD; PRT; 680 AA.
AC Q03692;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Collagen alpha 1(X) chain precursor.
GN COL10A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RX MEDLINE-92109659; PubMed-1764025;
RA Thomas J.T., Cresswell C.J., Rash B., Nicolai H., Jones T.,
RA Solomon E., Grant M.E., Boot-Handford R.P.;
RT "The human collagen X gene. Complete primary translated sequence and
RT chromosomal localization."
RL Biochem. J. 280:617-623(1991).
[2]
SEQUENCE FROM N.A.
RX MEDLINE-93012005; PubMed-1397333;
RA Reichenberger E., Beier F., Luvall P., Olsen B.R., von der Mark K.,
RA Bertling W.M.;
RT "Genomic organization and full-length cDNA sequence of human collagen
RT X."
RL FEBS Lett. 311:305-310(1992).
[3]
SEQUENCE FROM N.A.
RA Beier F., Lammi M.B., von der Mark K.;
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
RA Williams S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE OF 52-680 FROM N.A.
RX MEDLINE-92267014; PubMed-1597271;
RA Apte S.S., Seldin M.F., Hayashi M., Olsen B.R.;
RT "Cloning of the human and mouse type X collagen genes and mapping of
RT the mouse type X collagen gene to chromosome 10.";
```

```
RL Eur. J. Biochem. 206:217-224(1992).
RN [6]
RN SEQUENCE OF 561-666 FROM N.A.
RX MEDLINE-91243838; PubMed-2037056;
RA Apte S., Mattei M.G., Olsen B.R.;
RT "Cloning of human alpha 1(X) collagen DNA and localization of the
RT COL10A1 gene to the q21-q22 region of human chromosome 6."
RL FEBS Lett. 282:393-396(1991).
RN [7]
RN SEQUENCE OF 547-655 FROM N.A.
RX MEDLINE-92077285; PubMed-1743401;
RA Reichenberger E., Aigner T., von der Mark K., Stoeb H., Bertling W.;
RT "In situ hybridization studies on the expression of type X collagen
RT in fetal human cartilage."
RL Dev. Biol. 148:562-572(1991).
RN [8]
RN REVIEW ON VARIANTS.
RX MEDLINE-97255959; PubMed-9101290;
RA Kuivaniemi H., Tromp G., Prockop D.J.;
RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
RT associated collagen (type IX), and network-forming collagen (type X)
RT cause a spectrum of diseases of bone, cartilage, and blood vessels."
RL Hum. Mutat. 9:300-315(1997).
RN [9]
RN VARIANTS SMCD ASP-598 AND PRO-614.
RX MEDLINE-94136476; PubMed-8304336;
RA Wallis G.A., Rash B., Sweetman W.A., Thomas J.T., Super M., Evans G.,
RA Grant M.E., Boot-Handford R.P.;
RT "Amino acid substitutions of conserved residues in the
RT carboxyl-terminal domain of the alpha 1(X) chain of type X collagen
RT occur in two unrelated families with metaphyseal chondrodysplasia
RT type Schmid."
RL Am. J. Hum. Genet. 54:169-178(1994).
RN [10]
RN VARIANT SMCD ARG-591.
RX MEDLINE-94272470; PubMed-8004099;
RA McIntosh I., Abbott M.H., Warman M.L., Olsen B.R., Francmann C.A.;
RT "Additional mutations of type X collagen confirm COL10A1 as the
RT Schmid metaphyseal chondrodysplasia locus."
RL Hum. Mol. Genet. 3:303-307(1994).
RN [11]
RN VARIANT SMCD VAL-618.
RX MEDLINE-95181449; PubMed-7876225;
RA Chan D., Cole W.G., Rogers J.G., Bateman J.F.;
RT "Type X collagen multimer assembly in vitro is prevented by a Gly618
RT to Val mutation in the alpha 1(X) NC1 domain resulting in Schmid
RT metaphyseal chondrodysplasia."
RL J. Biol. Chem. 270:4558-4562(1995).
RN [12]
RN VARIANTS SMCD ARG-545; GLU-595; HIS-597; LYS-617; ARG-644 AND GLY-648.
RX MEDLINE-95311767; PubMed-7607655;
RA Bonaventure J., Chaminade F., Maroteaux P.;
RT "Mutations in three subdomains of the carboxy-terminal region of
RT collagen type X account for most of the Schmid metaphyseal
RT dysplasias."
RL Hum. Genet. 96:58-64(1995).
RN [13]
RN VARIANT SMCD PRO-600.
RX MEDLINE-96375754; PubMed-8782043;
RA Wallis G.A., Rash B., Sykes B., Bonaventure J., Maroteaux P.,
RA Zabel B., Wynne-Davies R., Grant M.E., Boot-Handford R.P.;
RT "Mutations within the gene encoding the alpha 1(X) chain of type X
RT collagen (COL10A1) cause metaphyseal chondrodysplasia type Schmid but
RT not several other forms of metaphyseal chondrodysplasia."
RL J. Med. Genet. 33:450-457(1996).
RN [14]
RN VARIANTS SMCD GLU-18 AND ARG-18.
RX MEDLINE-97220591; PubMed-9067753;
RA Ikegawa S., Nakamura K., Nagano A., Haga N., Nakamura Y.;
RT "Mutations in the N-terminal globular domain of the type X collagen
RT gene (COL10A1) in patients with Schmid metaphyseal
RT chondrodysplasia."
RL Hum. Mutat. 9:131-135(1997).
```

[15]
 RP VARIANTS SMD GLU-595.
 RX MEDLINE=99057503; PubMed=9837818;
 RA Ikegawa S., Nishimura G., Nagai T., Hasegawa T., Ohashi H.,
 RA Nakamura Y.;
 RT "Mutation of the type X collagen gene 'COL10A1' causes
 RT spondylometaphyseal dysplasia.";
 RL Am. J. Hum. Genet. 63:1659-1662(1998).
 [16]
 RP VARIANT SMD CYS-597.
 RX MEDLINE=99069781; PubMed=9852679;
 RA Sawai H., Ida A., Nakata Y., Koyama K.;
 RT "Novel missense mutation resulting in the substitution of tyrosine by
 RT cysteine at codon 597 of the type X collagen gene associated with
 RT Schmid metaphyseal chondrodysplasia.";
 RL J. Hum. Genet. 43:259-261(1998).
 CC -1- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTHROPHIC
 CC CHONDROCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
 CC MINERALIZATION ZONES OF HYALINE CARTILAGE.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1- DISEASE: DEFECTS IN COL10A1 ARE THE CAUSE OF SCHMID TYPE
 CC METAPHYSEAL CHONDRODYSPLASIA (SMCD), A DOMINANTLY INHERITED
 CC DISORDER OF THE OSSEOUS SKELETON. THE CARDINAL FEATURES OF THE
 CC PHENOTYPE ARE MILD SHORT STATURE, COXA VARA AND A WADDLING GAIT.
 CC RADIOGRAPHY USUALLY SHOWS SCLEROSIS OF THE RIBS, FLARING OF THE
 CC METAPHYSES, AND A WIDE IRREGULAR GROWTH PLATE, ESPECIALLY OF THE
 CC KNEES.
 CC -1- DISEASE: DEFECTS IN COL10A1 ARE THE CAUSE OF SPONDYLOMETAPHYSEAL
 CC DYSPLASIA (SMD). SMD COMPRISES A HETEROGENEOUS GROUP OF HERITABLE
 CC SKELETAL DYSPLASIAS CHARACTERIZED BY MODIFICATIONS OF THE
 CC VERTEBRAL BODIES OF THE SPINE AND METAPHYSES OF THE TUBULAR BONES.
 CC -1- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
 CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; X60382; CAA42933.1; -;
 DR EMBL; X65120; CAA46236.1; -;
 DR EMBL; X98568; CAA67178.1; -;
 DR EMBL; AL121963; CAB87590.1; -;
 DR EMBL; S68531; AAC60815.1; -;
 DR EMBL; X58879; CAA41686.1; -;
 DR EMBL; M74050; AAA61221.1; -;
 DR EMBL; X72579; CAA51170.1; -;
 DR EMBL; X72580; CAA51170.1; JOINED.
 DR PIR; S15826; S15826.
 DR PIR; S30086; S30086.
 DR PIR; A43901; A43901.
 DR PIR; S18249; S18249.
 DR PIR; S21856; S21856.
 DR PIR; S26396; S26396.
 DR Genew; HGNC:2185; COL10A1.
 DR MIM; 120110; -;
 DR MIM; 156500; -;
 DR MIM; 184250; -;
 DR InterPro; IPR001073; C1Q.
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF00386; C1Q; 1.
 DR Pfam; PF01391; Collagen; 8.
 DR PRINTS; PR00007; COMPLEMENTC1Q.
 DR SMART; SM00110; C1Q; 1.
 DR PROSITE; PS01113; C1Q; 1.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Cartilage; Collagen; Signal; Disease mutation; Polymorphism.
 FT SIGNAL 1 18 POTENTIAL.

FT CHAIN 19 680 COLLAGEN ALPHA 1(X) CHAIN.
 FT DOMAIN 19 56 NONHELICAL REGION (NC2).
 FT DOMAIN 57 519 TRIPLE-HELICAL REGION
 FT DOMAIN 520 680 NONHELICAL REGION (NC1).
 FT DOMAIN 545 680 C1Q.
 FT VARIANT 18 18 G -> E (IN SMCD).
 FT VARIANT 18 18 /FTID=VAR_001838.
 FT VARIANT 18 18 G -> R (IN SMCD).
 FT VARIANT 18 18 /FTID=VAR_001839.
 FT VARIANT 545 545 G -> R.
 FT VARIANT 591 591 /FTID=VAR_001840.
 FT VARIANT 591 591 C -> R (IN SMCD).
 FT VARIANT 591 591 /FTID=VAR_001841.

Query Match 27.7%; Score 362.5; DB 1; Length 680;
 Best Local Similarity 36.8%; Pred. No. 9.7e-17;
 Matches 86; Conservative 33; Mismatches 78; Indels 37; Gaps 9;
 QY 34 PHGPRG-----PGDGAPASVPPFPAGKGVGRGKAGLRGPPGPPGPPGPPGPPGPPG 89
 DB 453 PIGPPGIPGFPFGSKGDGGS--PGPPGAG-IATKGLNGTGPDPGPPGPPGPPGPPGPPG 509
 QY 90 PGPPGPPG-----GVAPAGVVPRIAFYAGLR-RPHEGYE----- 125
 DB 510 PGPPGP-PGQAVMPEGFIKAGQPSLGTPLVSANQGVTPGMPVSFTILSKAYPAIGTP 568
 QY 126 LRFDDVNTVNCVNEAAASGKETCPMPGVYFYAYHVMRGDGTSMWADLMKNGQVRSAL 185
 DB 569 IPFKILNROQHYPRTGIFTQIPGIYFSYHVHV---GTHVWVGLYNGTPTWMTY 625
 QY 186 AOADQONDYASNSVILHDVGEVFIKLDGKGVHGNNTNY--STFSGFIYP 237 -
 DB 626 DEYTKGYLDQASGSAIIDLTENDQVWLQLPNAESNGLYSSEYVHSSFSGLVAP 679

RESULT 9

CAIA_CHICK STANDARD; PRT; 674 AA.
 AC P08125;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Collagen alpha 1(X) chain precursor.
 GN COL10A1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE OF 48-674 FROM N.A., AND SEQUENCE OF 103-117 AND 453-466.
 RX MEDLINE=86168227; PubMed=3082876;
 RA Nimomiya Y., Gordon M., van der Rest M., Schmid T., Linsemayer T.,
 RA Olsen B.R.;
 RT "The developmentally regulated type X collagen gene contains a long
 RT open reading frame without introns.";
 RL J. Biol. Chem. 261:5041-5050(1986).
 [2]
 RP SEQUENCE OF 1-75 FROM N.A.
 RX MEDLINE=89054019; PubMed=2461368;
 RA Luvall P., Nimomiya Y., Rosenblum N.D., Olsen B.R.;
 RT "The type X collagen gene. Introns sequences split the 5'-untranslated
 RT region and separate the coding regions for the non-collagenous amino-
 RT terminal and triple-helical domains.";
 RL J. Biol. Chem. 263:18378-18385(1988).
 [3]
 RP REVISTONS TO C-TERMINUS.
 RX MEDLINE=89380199; PubMed=2476437;
 RA Yamaguchi N., Benya P.D., van der Rest M., Nimomiya Y.;
 RT "The cloning and sequencing of alpha 1(VIII) collagen cDNAs
 RT demonstrate that type VIII collagen is a short chain collagen and
 RT contains triple-helical and carboxyl-terminal non-triple-helical
 RT domains similar to those of type X collagen.";

615 VMTYDEYIKYLDQASGAVIDLTDENVMLQFLPAGNSGLYSPYVHSSFGSLVAP 673

RESULT 11

CLQB_HUMAN
ID CLQB_HUMAN STANDARD; PRT; 251 AA.
AC P02746; Q96H17;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Complement Clq subcomponent, B chain precursor.
GN ClQB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=86076906; PubMed=3000358;
RA Reid K.B.M.;
RT "Molecular cloning and characterization of the complementary DNA and
RT gene coding for the B-chain of subcomponent Clq of the human
RT complement system.";
RL Biochem. J. 231:729-735(1985).
RN [2]
RP SEQUENCE FROM N.A.
RC Tissue=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 26-133.
RX MEDLINE=80020137; PubMed=486087;
RA Reid K.B.M.;
RT "Complete amino acid sequences of the three collagen-like regions
RT present in subcomponent Clq of the first component of human
RT complement.";
RL Biochem. J. 179:367-371(1979).
RN [4]
RP SEQUENCE OF 26-193.
RX MEDLINE=79041552; PubMed=708376;
RA Reid K.B.M., Thompson E.O.P.;
RT "Amino acid sequence of the N-terminal 108 amino acid residues of the
RT B chain of subcomponent Clq of the first component of human
RT complement.";
RL Biochem. J. 173:863-868(1978).
RN [5]
RP SEQUENCE OF 134-251.
RX MEDLINE=82283890; PubMed=6981411;
RA Reid K.B.M., Gagnon J., Frampton J.;
RT "Completion of the amino acid sequences of the A and B chains of
RT subcomponent Clq of the first component of human complement.";
RL Biochem. J. 203:559-569(1982).
RN [6]
RP SEQUENCE OF 224-251 FROM N.A.
RC Tissue=Liver;
RX MEDLINE=85038855; PubMed=6208566;
RA Reid K.B.M., Bentley D.R., Wood K.J.;
RT "Cloning and characterization of the complementary DNA for the B
RT chain of normal human serum Clq.";
RL Philos. Trans. R. Soc. Lond., B, Biol. Sci. 306:345-354(1984).
RN [7]
RP REVIEW OF ClQ DEFICIENCY.
RX MEDLINE=98450587; PubMed=9777412;
RA Petry F.;
RT "Molecular basis of hereditary Clq deficiency.";
RL Immunobiology 199:286-294(1998).
CC -!- FUNCTION: ClQ ASSOCIATES WITH THE PROENZYME ClR AND ClS TO YIELD
CC Cl1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
CC COLLAGEN-LIKE REGIONS OF ClQ INTERACT WITH THE CA(2+)-DEPENDENT
CC ClR(2)ClS(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF Cl
CC TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF ClQ WITH THE
CC FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.

CC -!- SUBUNIT: Cl1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF ClQ, R
CC AND S IN THE MOLAR RATION OF 1:2:2. ClQ SUBCOMPONENT IS COMPOSED
CC OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE
CC A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF
CC THE C CHAIN.
CC -!- PTM: O-LINKED GLYCANS CONSIST OF GLC-GAL DISACCHARIDES.
CC -!- DISEASE: DEFECTS IN ClQ ARE A CAUSE OF ClQ DEFICIENCY. IT IS A
CC RARE GENETIC DISORDER WHICH IS ASSOCIATED WITH RECURRENT
CC INFECTIONS AND A HIGH PREVALENCE OF LUPUS ERYTHEMATOSUS-LIKE
CC SYMPTOMS. IT IS CHARACTERIZED BY A LOSS OF ACTIVATION OF THE
CC COMPLEMENT CLASSICAL PATHWAY.
CC -!- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 ClQ DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X03084; CAA26880.1;
DR EMBL: BC008983; AHA08983.1; ALT_INIT.
DR EMBL: M36278; AAC41692.1; -;
DR PIR: A03206; ClHUQB.
DR PIR: B23422; B23422.
DR Genew: HGNC:1242; ClQB.
DR MIM: 120570;
DR InterPro: IPR001073; Clq;
DR InterPro: IPR000087; Collagen.
DR Pfam: PF00386; Clq; 1.
DR Pfam: PF01391; Collagen; 1.
DR PRINTS: PR00007; COMPLEMENTClQ.
DR SMART: SM00110; ClQ; 1.
DR PROSITE: PS01113; Clq; 1.
KW Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
KW Repeat; Signal; Disease mutation.
FT SIGNAL 1 25
FT CHAIN 26 251
FT DOMAIN 29 112
FT DOMAIN 113 251
FT MOD_RES 26 26 PYRROLIDONE CARBOXYLIC ACID.
FT DISULFID 29 29 INTERCHAIN (WITH C-26 IN CHAIN A).
FT MOD_RES 33 33 HYDROXYLATION.
FT MOD_RES 36 36 HYDROXYLATION.
FT MOD_RES 39 39 HYDROXYLATION.
FT MOD_RES 42 42 HYDROXYLATION.
FT MOD_RES 51 51 HYDROXYLATION.
FT MOD_RES 54 54 HYDROXYLATION.
FT MOD_RES 57 57 HYDROXYLATION.
FT CARBOHYD 57 57 O-LINKED (GAL. .).
FT MOD_RES 60 60 HYDROXYLATION.
FT CARBOHYD 60 60 O-LINKED (GAL. .).
FT MOD_RES 63 63 HYDROXYLATION.
FT MOD_RES 75 75 HYDROXYLATION.
FT MOD_RES 81 81 HYDROXYLATION.
FT MOD_RES 84 84 HYDROXYLATION.
FT MOD_RES 90 90 HYDROXYLATION.
FT MOD_RES 96 96 HYDROXYLATION.
FT CARBOHYD 96 96 O-LINKED (GAL. .).
FT MOD_RES 99 99 HYDROXYLATION.
FT MOD_RES 102 102 HYDROXYLATION.
FT MOD_RES 105 105 HYDROXYLATION.
FT MOD_RES 108 108 HYDROXYLATION.
FT CARBOHYD 108 108 O-LINKED (GAL. .).
FT VARIANT 40 40 G -> D (IN ClQ DEFICIENCY).
FT FTIG=VAR_008541.
FT CONFLICT 26 26 Q -> E (IN REF. 3).
FT CONFLICT 83 83 N -> D (IN REF. 3).
FT CONFLICT 98 98 G -> P (IN REF. 3 AND 4).
FT SEQUENCE 251 AA; 26459 MW; 78C5752E267A0EF7 CRC64;
SQ

KW	Cell adhesion; Collagen.					
FT	NON_TER	1	1			
FT	DOMAIN	<1	11	NONHELICAL REGION (NC2).		
FT	DOMAIN	12	468	TRIPLE-HELICAL REGION.		
FT	DOMAIN	469	635	NONHELICAL REGION (NC1).		
FT	DOMAIN	500	635	C1Q.		
SQ	SEQUENCE	635 AA;	60527 MW;	22A261164754F771 CRC64;		
Query Match 26.3%; Score 345; DB 1; Length 635;						
Best Local Similarity 33.0%; Pred. No. 1.2e+15;						
Matches	96;	Conservative	32;	Mismatches	67;	Indels 96; Gaps 13;
QY	35	HGRPGP-GPDGAPA-SVPPFPFGA-----KGVEVRGRKAGLRGP-----	71			
Db	352	HGGPPGTGPKGEFTGCRPGCGPVACALGQGLDGLPGQPLGRPSGIPGLQGAPGPIGP	411			
QY	72	-----PGPP-----GPRGPGEPGRCPGPCPGC-PGPGGVAPAAGYV	107			
Db	412	QLPLGGLKGFELPGPGEGRAGEPGTAGRPPGPVPSPGIGTGPPLGPPGP-APGA	470			
QY	108	PRIAFYAGLRRPHGEYE-----	129			
Db	471	TGI---AGLHLPNGVGEGAVLGKKGKPQFGLGELSAAHTPAFTAVLTSLPASGMPVK	527			
QY	130	DVVTVNNGNYAEASGKETCPMGQVFYFAHYVLMRGGDGTSMMADLMKNQVRSATAQDA	189			
		: : : : : : : : : : : : : : :				
Db	528	RTLYNGSHSGYNPATGIFTCPVGCVYYFYAHVHVHK---GTNVVALYKN-NVPATYTYDEY	583			
QY	190	DQNY-DYANSVTILHLVDGDEVFIKLDDGKVHGCHNTNKY--STFSFIIYP	237			
		: : : : : : : : : : : : : : :				
Db	584	KKGYLDOASGGAYLRLRPNDQVWVQIPSDQANGLYSTETIHSSFSGLFLCP	634			
RESULT 13						
COT2_HUMAN STANDARD; PRT; 285 AA.						
ID	COT2_HUMAN					
AC	Q9BXJ5;					
DT	15-JUN-2002 (Rel. 41, Created)					
DT	15-JUN-2002 (Rel. 41, Last sequence update)					
DT	15-JUN-2002 (Rel. 41, Last annotation update)					
DE	Complement-clq tumor necrosis factor-related protein 2 precursor.					
DE	C1QTNF2 OR CTRP2.					
GN	Homo sapiens (Human).					
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
OX	NCBI_TaxID=9606;					
RN	[1]					
RA	SEQUENCE FROM N.A.					
RA	Piddington C.S.; Bishop P.;					
RT	"Homo sapiens complement-clq tumor necrosis factor-related protein.";					
RL	Submitted (DEC-2000) to the EMBL/GenBank/DDBJ databases.					
RN	[2]					
RA	SEQUENCE FROM N.A.					
RC	TISSUE=Muscle;					
RA	Straussberg R.;					
RL	Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.					
CC	- - SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.					
CC	- - SIMILARITY: CONTAINS 1 C1Q DOMAIN.					
CC						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/					
CC	or send an email to license@isb-sib.ch).					
DR	EMBL; AF329836; AAK17960.1; -					
DR	EMBL; BC011699; AAH11699.1; -					
DR	Gene; HGNC:14325; C1QTNF2.					
DR	InterPro; IPR001073; Clq.					
DR	InterPro; IPR000087; Collagen.					
DR	Pfam; PF00386; C1q; 1.					

RN	PFam: PF01391; Collagen; 2.
RP	PRINTS; PRO0007; COMPLEMNTCIQ.
RX	SMART; SM00110; C1Q; 1.
RA	PROSITE; PS01113; C1Q; 1.
DR	Collagen; Signal.
KW	SIGNAL.
FT	CHAIN
FT	16 285
FT	POTENTIAL.
FT	COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT	RELATED PROTEIN 2.
FT	COLLAGEN-LIKE.
FT	C1Q.
FT	DOMAIN
FT	40 141
FT	DOMAIN
FT	143 285
FT	SEQUENCE
FT	285 AA; 29952 MW; 7E31FF9868DAEDFA CRC64;
FT	Query Match
FT	Best Local Similarity 25.0%; Score 328; DB 1; Length 285;
FT	Matches 86; Conservative 37; Mismatches 83; Indels 50; Gaps 10;
FT	Matches
FT	26 GRCMVCDPHGPRG-PGPDGAPA-----SVPPPPGA 56
DB	30 GSPQLVSLPGPGPPGGPAGPSGMGRMGFPKGDQDGHDRSDSGEGPPGRTGN 89
QY	57 KGEVRRKGAGLRGPGRGPRGPRGPPGPPGPG-----PGGVPAAGIYP 108
DB	90 RCKPGPKKAGAIGRAGRPGKVGNGTPGKHGTGPKGPKGKEGLPCSCGSHT- 148
QY	109 RTAFVAGLR--PHEGYEVRLRDVVTVNGVAYAAASGKTCTCPMPGVYFYFAYHVMRGSD 166
DB	149 KSASFVAVTKSYPRERLPI-KFDKILMEGHYNASSGKFVCVPVIYYTYDTITLA--- 204
QY	167 GTSMADLMKNCOVRASAIAODAQ-NVDYASNVLHLVDGFVKIKLOGGKVHGNTN 225
DB	205 NKHLAIGHVNGQYRIPTF--DANTGNHDVASGSTILALKQGDEVWLQIFYSEQNGLFYD 262
QY	226 KY---STFGSLIYPD 238
DB	263 PYWDSLTGFLIYAD 278
RESULT 14	
C1QC_HUMAN	STANDARD; PRT; 245 AA.
ID	A02747; Q96DL2; Q96HU5;
AC	21-JUL-1986 (Rel. 01, Created)
DT	15-JUN-2002 (Rel. 41, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Complement Clq subcomponent, C chain precursor.
DN	C1Q OR C1QC.
GC	Homo sapiens (Human).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
ON	[1]
RE	SEQUENCE FROM N.A.
RC	TISSUE=Monoocytes;
RP	MEDLINE=91174759; PubMed=1706597;
RX	Sellar G.C., Blake D.J., Reid K.B.M.;
RA	"Characterization and organization of the genes encoding the A-, B-
RT	and C-chains of human complement subcomponent Clq. The complete
RT	derived amino acid sequence of human Clq.";
RT	Biochem. J. 274:481-490(1991).
RN	[2]
RN	SEQUENCE FROM N.A.
RP	TISSUE=Cerebellum;
RC	Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y.,
RA	Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
RA	Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
RA	Kawakami B., Nagai K., Isogai T., Sugano S.;
RT	"NEO human cDNA sequencing project."
RL	Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
RN	[3]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Brain;
RA	Straussberg R.;
RL	Submitted (May-2001) to the EMBL/GenBank/DBJ databases.

[4] SEQUENCE OF 29-122.
MEDLINE=80020137; PubMed=486087;
Reid K.B.M.;
"Complete amino acid sequences of the three collagen-like regions
present in subcomponent Clq of the first component of human
complement.";
Biochem. J. 179:367-371(1979).
[5]
REVIEW OF C1Q DEFICIENCY.
MEDLINE=98450587; PubMed=9777412;
Petty F.;
"Molecular basis of hereditary C1q deficiency.";
Immunobiology 199:286-294(1998).
C1- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
C1- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
AND S IN THE MOLAR RATIO OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED
OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE
A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF
THE C CHAIN.
C1- PTM: O-LINKED GLYCANS CONSIST OF GLC-GAL DISACCHARIDES.
C1- DISEASE: DEFECTS IN C1Q ARE A CAUSE OF C1Q DEFICIENCY. IT IS A
RARE GENETIC DISORDER WHICH IS ASSOCIATED WITH RECURRENT
INFECTIONS AND A HIGH PREVALENCE OF LUPUS ERYTHEMATOSUS-LIKE
SYMPTOMS. IT IS CHARACTERIZED BY A LOSS OF ACTIVATION OF THE
COMPLEMENT CLASSICAL PATHWAY.
C1- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
C1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; AK057792; BAB71575.1; -.
EMBL; BC009016; AAA09016.1; -.
PIR; A03207; C1HUQC.
PIR; SI4351; SI4351.
Gene; HGNC:1245; C1QG.
MIM; 120575; -.
InterPro; IPR001073; Clq.
InterPro; IPR000087; Collagen.
Pfam; PF00386; Clq; 1.
Pfam; PF01391; Collagen; 1.
PRINTS; PRO0007; COMPLEMNTCIQ.
SMART; SM00110; C1Q; 1.
PROSITE; PS01113; C1Q; 1.
Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
Repeat; Signal; Disease mutation.
SIGNAL 1 28
CHAIN 29 245 COMPLEMENT C1Q SUBCOMPONENT, C CHAIN.
DOMAIN 31 112 COLLAGEN-LIKE.
DOMAIN 113 245 C1Q.
DISULFID 32 32 INTERCHAIN (WITH OTHER C CHAIN).
MOD_RES 36 36 HYDROXYLATION.
MOD_RES 39 39 HYDROXYLATION.
MOD_RES 42 42 HYDROXYLATION.
MOD_RES 45 45 HYDROXYLATION.
MOD_RES 54 54 HYDROXYLATION.
MOD_RES 57 57 HYDROXYLATION.
MOD_RES 63 63 HYDROXYLATION.
MOD_RES 66 66 HYDROXYLATION.
MOD_RES 71 71 HYDROXYLATION.
MOD_RES 75 75 HYDROXYLATION.
CARBOHYD 75 75 O-LINKED (GAL.).

RN	PFam: PF01391; Collagen; 2.
RP	PRINTS; PRO0007; COMPLEMNTC1Q.
RX	SMART; SM00110; C1Q; 1.
RA	PROSITE; PS01113; C1Q; 1.
DR	Collagen; Signal.
KW	SIGNAL.
FT	CHAIN
FT	16 285
FT	POTENTIAL.
FT	COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT	RELATED PROTEIN 2.
FT	COLLAGEN-LIKE.
FT	C1Q.
FT	DOMAIN
FT	40 141
FT	DOMAIN
FT	143 285
FT	SEQUENCE
FT	285 AA; 29952 MW; 7E31FF9868DAEDFA CRC64;
FT	Query Match
FT	Best Local Similarity 25.0%; Score 328; DB 1; Length 285;
FT	Matches 86; Conservative 37; Mismatches 83; Indels 50; Gaps 10;
FT	Matches
FT	26 GRCMVCDPHGPRG-PGPDGAPA-----SYVPPPPGA 56
DB	30 GSPQLVSLPGPGPPGGPAGPSGMGRMGFPKGDQDGHDRSDSEEGPPGRTGN 89
QY	57 KGEVRRKGAGLRGPGRGPRGPRGPPGPPGPG-----PGGVAPAAGYVP 108
DB	90 RCKPGPKKAGAIGRAGRPGKVGNGTPGKHGTGPKGPKGKEGLPCSGSGHT- 148
QY	109 RTAFVAGLRR--PHEGYEVRLRDVVTVNVGNAYRAASGKTFCMPMGVYFYFAYHVMRGSD 166
DB	149 KSAFSVAVTYSKYPRERLP1-KFDKILMEGHYNASSGKFVCVGPVIYYTYDTITLA--- 204
QY	167 GTSMADLMKNCOVRASAIAODAQ-NVDYASNVLHLVDGFVKIKLOGGKVHGNTN 225
DB	205 NKHLAIGLVHNGQYRIPTF--DANTGNHDVASGSTILALKQGDEVWLQIFYSEQNGLFYD 262
QY	226 KY---STFGSLIYPD 238
DB	263 PYWDSLTGFLIYAD 278
RESULT 14	
C1QC_HUMAN	STANDARD; PRT; 245 AA.
ID	P02747; Q96DL2; Q96HU5;
AC	21-JUL-1986 (Rel. 01, Created)
DT	15-JUN-2002 (Rel. 41, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Complement C1q subcomponent, C chain precursor.
DN	C1Q OR C1QC.
GC	Homo sapiens (Human).
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
ON	[1]
RE	SEQUENCE FROM N.A.
RC	TISSUE=Monoocytes;
RP	MEDLINE=91174759; PubMed=1706597;
RX	Sellar G.C., Blake D.J., Reid K.B.M.;
RA	"Characterization and organization of the genes encoding the A-, B-
RA	and C-chains of human complement subcomponent C1q. The complete
RT	derived amino acid sequence of human C1q.";
RT	Biochem. J. 274:481-490(1991).
RN	[2]
RN	SEQUENCE FROM N.A.
RP	TISSUE=Cerebellum;
RC	Nishi T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y.,
RA	Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
RA	Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
RA	Kawakami B., Nagai K., Isogai T., Sugano S.;
RA	"NEO human cDNA sequencing project."
RL	Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
RN	[3]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Brain;
RA	Straussberg R.;
RL	Submitted (May-2001) to the EMBL/GenBank/DBJ databases.

[4] SEQUENCE OF 29-122.
MEDLINE=80020137; PubMed=486087;
Reid K.B.M.;
"Complete amino acid sequences of the three collagen-like regions
present in subcomponent Clq of the first component of human
complement.";
Biochem. J. 179:367-371(1979).
[5]
REVIEW OF C1Q DEFICIENCY.
MEDLINE=98450587; PubMed=9777412;
Petty F.;
"Molecular basis of hereditary C1q deficiency.";
Immunobiology 199:286-294(1998).
C1- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1r AND C1s TO YIELD
COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
C1- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
AND S IN THE MOLAR RATIO OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED
OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE
A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF
THE C CHAIN.
C1- PTM: O-LINKED GLYCANS CONSIST OF GLC-GAL DISACCHARIDES.
C1- DISEASE: DEFECTS IN C1Q ARE A CAUSE OF C1Q DEFICIENCY. IT IS A
RARE GENETIC DISORDER WHICH IS ASSOCIATED WITH RECURRENT
INFECTIONS AND A HIGH PREVALENCE OF LUPUS ERYTHEMATOSUS-LIKE
SYMPTOMS. IT IS CHARACTERIZED BY A LOSS OF ACTIVATION OF THE
COMPLEMENT CLASSICAL PATHWAY.
C1- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
C1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```
FT MOD_RES 81 81 HYDROXYLATION.
FT MOD_RES 84 84 HYDROXYLATION.
FT CARBOHYD 84 84 O-LINKED (GAL. . .).
FT MOD_RES 93 93 HYDROXYLATION.
FT MOD_RES 96 96 HYDROXYLATION.
FT MOD_RES 99 99 HYDROXYLATION.
FT MOD_RES 105 105 HYDROXYLATION.
FT VARIANT 43 43 G -> R (IN C1Q DEFICIENCY).
FT 14 14 /FTID=VAR_008542.
FT CONFLICT 14 14 K -> R (IN REF. 2).
FT CONFLICT 23 23 P -> A (IN REF. 1).
FT CONFLICT 57 57 K -> P (IN REF. 4).
FT CONFLICT 66 66 P -> K (IN REF. 4).
FT CONFLICT 72 72 K -> P (IN REF. 4).
FT CONFLICT 84 84 P -> K (IN REF. 4).
FT CONFLICT 87 87 N -> D (IN REF. 4).
FT CONFLICT 90 90 M -> N (IN REF. 4).
FT CONFLICT 215 215 E -> G (IN REF. 2).
SQ SEQUENCE 245 AA; 25774 MW; FAI7I7EB7ABFC12 CRC64;
Query Match 24.9%; Score 327; DB 1; Length 245;
Best Local Similarity 34.5%; Pred. No. 7.2e-15;
Matches 86; Conservative 36; Mismatches 97; Indels 30; Gaps 9;
QY 2 VLLLVVAIPLLVHSSRGPAHYEMLGRCRMVCDPHGRGPGDG-----APASVPPPP--P 54
Db 15 LLLLLLLLPL-----RQANTGCGYIGPM---PGLPGAPGKDGDLPGPKGPGIPAI 66
QY 55 GAKGEVGRGKAGLRGPPGPGPRGPGPRGPGPGPG-PGPGGVAAGVVPRI-AF 112
Db 67 GIRGPKGQGEPCGLGPKGNGKMGPPGPGPGVPGMPGIPGEPGEG-----RYKQFQSV 121
QY 113 YAGLRPRPE---GYEVLRFDDVVTNNGVNAEASGKFTCPMPGCVFFAYHVLMRG 169
Db 122 FTVTQTHPPAPNLSIRENAVLNPGQDYDTSTGKFTCKVPGLYEYVYH-----ASH 176
QY 170 MWADLMKNGQVRASATAQADQNDYASNSVILHLDYGVFVKLDGKGVGHGNTWKYST 229
Db 177 NLGVLLSRGVKVVYVFCGHTSKTNQVNSGVLRLQGVGEVWLAVNDYDMVGIOGSDSV 236
QY 230 FSGFIYPD 238
Db 237 FSGFLLFPD 245
RESULT 15
APM1_HUMAN STANDARD; PRT; 244 AA.
AC Q15848;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Adiponectin precursor (30 kDa adipocyte complement-related protein)
DE (ACRP30) (Adipose most abundant gene transcript 1) (apM-1) (Gelatin-
DE binding protein).
GN APM1 OR ACRP30 OR GBP28.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipose tissue;
RX MEDLINE=96224171; PubMed=8619847;
RA Maeda K., Okubo K., Shimomura I., Funahashi T., Matsuzawa Y.,
RA Matsubara K.;
RT "cDNA cloning and expression of a novel adipose specific collagen-like
RT factor, apM1 (Adipose Most abundant Gene transcript 1).";
RL Biochem. Biophys. Res. Commun. 221:286-289(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99196984; PubMed=10095105;
RA Saito K., Tobe T., Minoshima S., Asakawa S., Sumiya J., Yoda M.,
```

```
RA Nakano Y., Shimizu N., Tomita M.;
RT "Organization of the gene for gelatin-binding protein (GBP28).";
RL Gene 229:67-73(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99333693; PubMed=10403784;
RA Schaeffler A., Orso E., Palitzsch K.D., Buechler C., Drobnik W.,
RA Fuerst A., Schoelmerich J., Schmitz G.;
RT "The human apM-1, an adipocyte-specific gene linked to the family of
RT TNF's and to genes expressed in activated T cells, is mapped to
RT chromosome 1q21.3-q23, a susceptibility locus identified for familial
RL combined hyperlipidemia (FCH).";
RN [4]
RP Biochem. Biophys. Res. Commun. 260:416-425(1999).
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=20417747; PubMed=10961870;
RA Yokota T., Oritani K., Takahashi I., Ishikawa J., Matsuyama A.,
RA Ouchi N., Kihara S., Funahashi T., Tenner A.J., Tomiyama Y.,
RA Matsuzawa Y.;
RT "Adiponectin, a new member of the family of soluble defense collagens,
RT negatively regulates the growth of myelomonocytic progenitors and the
RT functions of macrophages.";
RL Blood 96:1723-1732(2000).
RN [5]
RP CHARACTERIZATION.
RX MEDLINE=20440368; PubMed=10982546;
RA Ouchi N., Kihara S., Arita Y., Okamoto Y., Maeda K., Kuriyama H.,
RA Hotta K., Nishida M., Takahashi M., Muraguchi M., Ohmoto Y.,
RA Nakamura T., Yamashita S., Funahashi T., Matsuzawa Y.;
RT "Adiponectin, an adipocyte-derived plasma protein, inhibits
RT endothelial NF-kappaB signaling through a cAMP-dependent pathway.";
RL Circulation 102:1296-1301(2000).
RN [6]
RP FUNCTION.
RX MEDLINE=21372498; PubMed=11479637;
RA Yamauchi T., Kamon J., Waki H., Terauchi Y., Kubota N., Hara K.,
RA Mori Y., Ide Y., Murakami K., Tsuboyama-Kasaoka N., Ezaki O.,
RA Akanuma Y., Gavrilova O., Vinson C., Reitman M.L., Kagechika H.,
RA Shudo K., Yoda M., Nakano Y., Tobe K., Nagai R., Kimura S., Tomita M.,
RA Froguel P., Kadowaki T.;
RT "The fat-derived hormone adiponectin reverses insulin resistance
RT associated with both lipodystrophy and obesity.";
RL Nat. Med. 7:941-946(2001).
RN [7]
RP VARIANT ADIPONECTIN DEFICIENCY CYS-112.
RX MEDLINE=20378830; PubMed=10918532;
RA Takahashi M., Arita Y., Yamagata K., Matsukawa Y., Okutomi K.,
RA Horie M., Shimomura I., Hotta K., Kuriyama H., Kihara S., Nakamura T.,
RA Yamashita S., Funahashi T., Matsuzawa Y.;
RT "Genomic structure and mutations in adipose-specific gene,
RT adiponectin.";
RL Int. J. Obes. Relat. Metab. Disord. 24:861-868(2000).
RN [8]
RP VARIANTS ARG-84; MET-117; THR-164; SER-221 AND PRO-241.
RX MEDLINE=21671103; PubMed=11812766;
RA Hara K., Boutin P., Mori Y., Tobe K., Dina C., Yasuda K., Yamauchi T.,
RA Otabe S., Okada T., Eto K., Kadowaki H., Hagura R., Akanuma Y.,
RA Yazaki Y., Nagai R., Taniyama M., Matsubara K., Yoda M., Nakano Y.,
RA Kimura S., Tomita M., Kimura S., Ito C., Froguel P., Kadowaki T.;
RT "Genetic variation in the gene encoding adiponectin is associated with
RT an increased risk of type 2 diabetes in the Japanese population.";
RL Diabetes 51:536-540(2002).
RN [9]
RP FUNCTION: IMPORTANT NEGATIVE REGULATOR IN HEMATOPOIESIS AND IMMUNE
CC SYSTEMS; MAY BE INVOLVED IN ENDING INFLAMMATORY RESPONSES THROUGH
CC ITS INHIBITORY FUNCTIONS. INHIBITS ENDOTHELIAL NF-KAPPAB SIGNALING
CC THROUGH A CAMP-DEPENDENT PATHWAY. INHIBITS TNF-ALPHA-INDUCED
CC EXPRESSION OF ENDOTHELIAL ADHESION MOLECULES. INVOLVED IN THE
CC CONTROL OF FAT METABOLISM AND INSULIN SENSITIVITY.
CC -1- SUBUNIT: HOMODIGLIMER (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: SECRETED IN PLASMA.
CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN THE ADIPOSE TISSUE.
CC -1- DISEASE: Defects in APML are the cause of adiponectin deficiency,
CC resulting in very low concentration of plasma adiponectin.
```

CC Decreased adiponectin plasma levels are associated with obesity
CC insulin resistance, and diabetes type 2.
CC -!- PHARMACEUTICAL: Adiponectin might be used in the treatment of
CC diabetes type 2 and insulin resistance.
CC -!- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

DR EMBL; D45371; BAA08227.1; -
DR EMBL; AB012165; BAA86716.1; -
DR EMBL; AB012164; BAA86716.1; JOINED.
DR EMBL; AJ131460; CAB52413.1; -
DR EMBL; AJ131461; CAB52413.1; JOINED.
DR MIM; 605441; -
DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Hormone; Collagen; Signal; Repeat; Hydroxylation; Plasma;
KW Polymorphism; Disease mutation; Obesity; Diabetes mellitus.
FT SIGNAL 1 14
FT CHAIN 15 244
FT DOMAIN 42 107
FT DOMAIN 108 244
FT DISULFID 36 36
FT MOD_RES 44 44
FT MOD_RES 47 47
FT MOD_RES 53 53
FT MOD_RES 62 62
FT MOD_RES 71 71
FT MOD_RES 76 76
FT MOD_RES 86 86
FT MOD_RES 95 95
FT MOD_RES 104 104
FT VARIANT 84 84
FT VARIANT 112 112
FT VARIANT 117 117
FT VARIANT 164 164
FT VARIANT 221 221
FT VARIANT 241 241
SQ SEQUENCE 244 AA; 26414 MW; 64D8C6C1204B1018 CRC64;
R -> C (IN ADIPONECTIN DEFICIENT).
V -> M.
I -> T.
R -> S.
H -> P.
/FTid=VAR_013273.
/FTid=VAR_013274.
/FTid=VAR_013275.
/FTid=VAR_013276.
/FTid=VAR_013277.
/FTid=VAR_013278.

Query Match 24.3%; Score 318; DB 1; Length 244;
Best Local Similarity 34.3%; Pred. No. 2.7e-14;
Matches 92; Conservative 36; Mismatches 82; Indels 58; Gaps 12;
QY 3 LLLLVATPLLVHSSRGAHYEMGLGRCRMVCDPHGPRPG-----PDGA-----PASVPPFP 53
DB 1 MLLLGAVLLLLAL---PCHDQ-----ETTQGPVLLPLPKGACTGWMAGIPGHP 47
QY 54 -----PGAKGVGRGKAGLRGPPGPPGPPGPPGPPGPPGPPG-----PGPG 98
DB 48 GHNGAPGRDGRDTPGKGEKGDPLIGPKDIGETGVPGAEGPRGPGIQRKGEPEG 107
QY 99 GVAPAGVVPRIATYAGLRPHGEYEV-----LRFDDVTVNVGNAYEASGKFTCPMPGV 153
DB 108 -----AYVRSASFVGL-----ETYVTIPNMPIRFTKIFYNQNNHYDSTGKFCNIPGL 157

OY 154 YFFAYHVLMRGGDGTSMWADLMKNGQVRASATAQDADQNDYVANSVILHLVDVGDEVFK 213
DB 158 YFFAYHITVYMKD---VKVSLFKKDKKAMLFYDQYQENNVDQASGVLLHLEVGQVWLQ 214
QY 214 LDG-GKVHG--GNTNKYSTFGFIYYPD 238
DB 215 VYGEGRNGLYADNDNDSTFTGFLLYHD 242

Search completed: June 2, 2003, 16:45:59
Job time : 23 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query Match	Length	DB	ID	Description
1	327	24.9	294	4	US-09-188-930-294	Sequence 294, Appl
2	318	24.3	244	2	US-08-463-911-7	Sequence 7, Appl
3	318	24.3	244	4	US-09-140-804-3	Sequence 3, Appl
4	318	24.3	244	4	US-09-336-536-20	Sequence 20, Appl
5	318	24.3	244	4	US-09-530-423-1	Sequence 1, Appl
6	317.5	24.2	247	2	US-08-463-911-2	Sequence 2, Appl
7	316.5	24.1	231	4	US-09-530-423-2	Sequence 2, Appl
8	316	24.1	246	2	US-08-463-911-4	Sequence 4, Appl
9	314	24.0	245	4	US-09-140-804-4	Sequence 4, Appl
10	309.5	23.6	247	4	US-09-140-804-8	Sequence 8, Appl
11	309.5	23.6	247	4	US-09-118-408-3	Sequence 3, Appl
12	309.5	23.6	247	4	US-09-506-855-3	Sequence 3, Appl
13	300	22.9	228	4	US-09-336-536-4	Sequence 4, Appl
14	300	22.9	243	4	US-09-188-930-295	Sequence 295, Appl
15	300	22.9	243	4	US-09-140-804-2	Sequence 2, Appl
16	300	22.9	243	4	US-09-336-536-3	Sequence 3, Appl
17	299	22.8	228	4	US-09-336-536-11	Sequence 11, Appl
18	299	22.8	243	4	US-09-336-536-10	Sequence 10, Appl
19	268.5	20.5	215	4	US-09-140-804-5	Sequence 5, Appl
20	257	19.6	236	4	US-09-140-804-6	Sequence 6, Appl
21	254.5	19.4	185	2	US-08-463-911-3	Sequence 3, Appl
22	250	19.1	423	1	US-08-383-744-2	Sequence 2, Appl
23	250	19.1	423	2	US-08-999-336-2	Sequence 2, Appl
24	250	19.1	423	5	PCT-US96-01427-2	Sequence 2, Appl
25	235.5	18.0	330	1	US-08-642-253-32	Sequence 32, Appl
26	235.5	18.0	408	1	US-07-609-716-65	Sequence 65, Appl
27	235.5	18.0	408	4	US-08-475-411A-65	Sequence 65, Appl

; Sequence 7, Application US/08463911
; Patent No. 5869330
; GENERAL INFORMATION:
; APPLICANT: Scherer, Philipp E.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED
; TITLE OF INVENTION: EXCLUSIVELY IN ADIPOCYTES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,911
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WHI95-05
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 861-6240
; TELEFAX: (617) 861-9540
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-463-911-7

Query Match 24.3%; Score 318; DB 2; Length 244;
Best Local Similarity 34.3%; Pred. No. 6e-20;
Matches 92; Conservative 36; Mismatches 82; Indels 58; Gaps 12;
QY 3 LLLVAIPLLVHSSRGPAHYEMLGRCRMVCDPHGPRPG-----PDGA-----PASVPPPP 53
Db 1 MLLGAVLLALL--PGHDQ-----ETTTQGGVLLPLPKGACTGWMAGIPGHP 47
QY 54 -----PGAKEVGRCKAGLRGPPGPPGPPGPPGPPGPPG-----PGPG 98
Db 48 GHNGAPGRDRGDTGPEKGEKGDPLGPKDGETGVGAEGRGFPGIGRKGPEGEG 107
QY 99 GVAPAGYYPPIAIFYAGLRPHEGYEV-----LRFDVVTVNMGVYASNSVILHLVDGVEFTK 153
Db 158 YFAYHIVMRGDDGTSWADLMKNGQVRSATQADQNDYQENNVQASGVLLHLEVDQVWLQ 214
QY 214 LDG-GKVHG--GNTNKYSTFSGFIIPD 238
Db 215 VYGEGERNGLYADNDNDSTFTGFLLYHD 242

RESULT 3
US-09-140-804-3
; Sequence 3, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Humes, Jacqueline M.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-49

; CURRENT APPLICATION NUMBER: US/09/140,804
; CURRENT FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: 60/056,983
; EARLIER FILING DATE: 1997-08-26
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-140-804-3

Query Match 24.3%; Score 318; DB 4; Length 244;
Best Local Similarity 34.3%; Pred. No. 6e-20;
Matches 92; Conservative 36; Mismatches 82; Indels 58; Gaps 12;
QY 3 LLLVAIPLLVHSSRGPAHYEMLGRCRMVCDPHGPRPG-----PDGA-----PASVPPPP 53
Db 1 MLLGAVLLALL--PGHDQ-----ETTTQGGVLLPLPKGACTGWMAGIPGHP 47
QY 54 -----PGAKEVGRCKAGLRGPPGPPGPPGPPGPPGPPG-----PGPG 98
Db 48 GHNGAPGRDRGDTGPEKGEKGDPLGPKDGETGVGAEGRGFPGIGRKGPEGEG 107
QY 99 GVAPAGYYPPIAIFYAGLRPHEGYEV-----LRFDVVTVNMGVYASNSVILHLVDGVEFTK 153
Db 108 -----AYVRSFAFVGL-----ETTYTIPNMPIRFTKIFYNQNHVDGSGTKFHCNIPGL 157
QY 154 YFAYHIVMRGDDGTSWADLMKNGQVRSATQADQNDYQENNVQASGVLLHLEVDGVEFTK 213
Db 158 YFAYHITVYMKD---VKVSLFKKDKAMLTFTDYQENNVQASGVLLHLEVDQVWLQ 214
QY 214 LDG-GKVHG--GNTNKYSTFSGFIIPD 238
Db 215 VYGEGERNGLYADNDNDSTFTGFLLYHD 242

RESULT 4
US-09-336-536-20
; Sequence 20, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-336-536-20

Query Match 24.3%; Score 318; DB 4; Length 244;
Best Local Similarity 34.3%; Pred. No. 6e-20;
Matches 92; Conservative 36; Mismatches 82; Indels 58; Gaps 12;
QY 3 LLLVAIPLLVHSSRGPAHYEMLGRCRMVCDPHGPRPG-----PDGA-----PASVPPPP 53
Db 1 MLLGAVLLALL--PGHDQ-----ETTTQGGVLLPLPKGACTGWMAGIPGHP 47
QY 54 -----PGAKEVGRCKAGLRGPPGPPGPPGPPGPPGPPG-----PGPG 98
Db 48 GHNGAPGRDRGDTGPEKGEKGDPLGPKDGETGVGAEGRGFPGIGRKGPEGEG 107
QY 99 GVAPAGYYPPIAIFYAGLRPHEGYEV-----LRFDVVTVNMGVYASNSVILHLVDGVEFTK 153
Db 108 -----AYVRSFAFVGL-----ETTYTIPNMPIRFTKIFYNQNHVDGSGTKFHCNIPGL 157

QY 154 YFFAYHVLRCGDTSMWADLMKNGQVRASIAQADQNYDYASNSVILHLDVGDEVEIK 213
Db 158 YFFAYHITVYMKD---VKVSLFKKDKAMLFYDQYQENNVQASGVLLHLEVGDQVWLQ 214
QY 214 LDG-GKVHG--GNTNKYSTFSGFIIPD 238
Db 215 VYGEGERNGLYADNDNDSTFTGFLLYHD 242

RESULT 5
US-09-530-423-1
; Sequence 1, Application US/09530423
; Patent No. 6461821
; GENERAL INFORMATION:
; APPLICANT: Otsuka Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: Smooth muscle growth inhibitory composition, a
; TITLE OF INVENTION: diagnostic method for arteriosclerosis and a kit
; TITLE OF INVENTION: therefor
; FILE REFERENCE: P98-51
; CURRENT APPLICATION NUMBER: US/09/530,423
; CURRENT FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: JP H9-297569
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Abdominal fat tissue from myoma uteri
US-09-530-423-1

Query Match 24.3%; Score 318; DB 4; Length 244;
Best Local Similarity 34.3%; Pred. No. 6e-20;
Matches 92; Conservative 36; Mismatches 82; Indels 58; Gaps 12;

QY 3 LLLVAIPLLVHSSRGPAHYEMLGRCRMVCDPHGPGP-----PDGA-----PASVPPPP 53
Db 1 MLLGAVLLLLAL---PGHDQ-----ETTTQGGVLLPLPKGACTGMAGIPGHP 47

QY 54 -----PGAKGVGRGKAGLRGPPGPGPRGPGGPRGPPGPG-----PGPG 98
Db 48 GHNAGPRGDRGTGPEKGEKGLGIPGKDIETGVPGAEGRPGFGIOGRKEGPEG 107

QY 99 GVAPAAGYVPRIAFYAGLRPRPHEGYEY-----LRFDDVYTVNVGNAYEAASGKFTCPMPGV 153
Db 108 -----AVYRSASFVGL-----ETVVTIPNPIRFTKIFYNQNHVDGSKGFCNIPGL 157

QY 154 YFFAYHVLRCGDTSMWADLMKNGQVRASIAQADQNYDYASNSVILHLDVGDEVEIK 213
Db 158 YFFAYHITVYMKD---VKVSLFKKDKAMLFYDQYQENNVQASGVLLHLEVGDQVWLQ 214

QY 214 LDG-GKVHG--GNTNKYSTFSGFIIPD 238
Db 215 VYGEGERNGLYADNDNDSTFTGFLLYHD 242

RESULT 6
US-08-463-911-2
; Sequence 2, Application US/08463911
; Patent No. 5869330
; GENERAL INFORMATION:
; APPLICANT: Scherer, Philipp E.
; APPLICANT: Lodish, Harvey F.
; TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED
; TITLE OF INVENTION: EXCLUSIVELY IN ADIPOCYTES
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173

QY 154 YFFAYHVLRCGDTSMWADLMKNGQVRASIAQADQNYDYASNSVILHLDVGDEVEIK 213
Db 158 YFFAYHITVYMKD---VKVSLFKKDKAMLFYDQYQENNVQASGVLLHLEVGDQVWLQ 214
QY 214 LDG-GKVHG--GNTNKYSTFSGFIIPD 238
Db 215 VYGEGERNGLYADNDNDSTFTGFLLYHD 242

RESULT 7
US-09-530-423-2
; Sequence 2, Application US/09530423
; Patent No. 6461821
; GENERAL INFORMATION:
; APPLICANT: Otsuka Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: Smooth muscle growth inhibitory composition, a
; TITLE OF INVENTION: diagnostic method for arteriosclerosis and a kit
; TITLE OF INVENTION: therefor
; FILE REFERENCE: P98-51
; CURRENT APPLICATION NUMBER: US/09/530,423
; CURRENT FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: JP H9-297569
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Abdominal fat tissue from myoma uteri
US-09-530-423-2

Query Match 24.1%; Score 316.5; DB 4; Length 231;
Best Local Similarity 36.1%; Pred. No. 7.6e-20;

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,911
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WHI95-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-911-2

Query Match 24.2%; Score 317.5; DB 2; Length 247;
Best Local Similarity 34.7%; Pred. No. 6.7e-20;
Matches 90; Conservative 32; Mismatches 94; Indels 43; Gaps 10;

QY 3 LLLVAIPLLVHSSRGPAHYEMLGRCRMVCDPHG-----PRGPGDGPASVPPPP 53
Db 7 LLFLILP---SHAEDDVTTTELAPA-LVPPPKGTCAGMMAGIPGPHGHTPGR--DGR 61

QY 54 PGAKGEVGRGKAGLRG-----PGPPGPRGPGGPRGPPGPGGPGGVAPAA 104
Db 62 DGTPEKGEKGDAGLLGPKGETGDMGTGAEGPRGPGTGRKGEFG-----EA 110

QY 105 GYVPRIAFYAGLR-RPHEGYEVLRFDDVYTVNVGNAYEAASGKFTCPMPGVYFFAYHVL 163
Db 111 AMYRSASFVGLTRVTVNPIRFTKIFYNQNHVDGSKGFCNIPGLYFSYHITVY 170

QY 164 GDDGTSMWADLMKNGQVRASIAQADQNYDYASNSVILHLDVGDEVEIKLDGKVHG- 222
Db 171 MKD---VKVSLFKKDKAVLFTYDQYQENNVQASGVLLHLEVGDQVWLQVYGDGHNGL 227

QY 223 ---NTNKYSTFSGFIIPD 238
Db 228 YADNVND-STFTGFLLYHD 245

RESULT 7
US-09-530-423-2
; Sequence 2, Application US/09530423
; Patent No. 6461821
; GENERAL INFORMATION:
; APPLICANT: Otsuka Pharmaceutical Co., Ltd.
; TITLE OF INVENTION: Smooth muscle growth inhibitory composition, a
; TITLE OF INVENTION: diagnostic method for arteriosclerosis and a kit
; TITLE OF INVENTION: therefor
; FILE REFERENCE: P98-51
; CURRENT APPLICATION NUMBER: US/09/530,423
; CURRENT FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: JP H9-297569
; PRIOR FILING DATE: 1997-10-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Abdominal fat tissue from myoma uteri
US-09-530-423-2

Query Match 24.1%; Score 316.5; DB 4; Length 231;
Best Local Similarity 36.1%; Pred. No. 7.6e-20;

Matches 84; Conservative 32; Mismatches 72; Indels 45; Gaps 10;

```
QY 38 RGP-----PQCA-----PASVPPFP-----PGAKEVRRGKAGLRGPPGPPGPPGPE 83
Db 10 QGPGVLLPLPKACTGWAIGPCHGNGCAPGRDGTTPGKEGKGPGLGPKGDIGE 69
QY 84 PGRPGPPGPG-----PGPGVAPAGYVPRIFAYAGLRPHGEYEV-----LRF 128
Db 70 TCVPGAEGPRGPGIQRKKEGEG-----AYVRSAPFSVGL-----ETVVTIPNMPIRF 119
QY 129 DDVTVNNGNAYEAASGKTCMPGYYFFAYHVLMRGGDGTSMWADLMKNGQVRASIAQD 188
Db 120 TKIFYNQNHQDGTGKHCNTPGLYYFAYHITVTYMKD---VKVSLFKKDKKAMLFYDQY 176
QY 189 ADQNTDYASNSVILHLDVGEYFIKLDG-GKVGH--GNTNKYSTFSGFIYPD 238
Db 177 QENNVDAQSGVLLHLEVDQWLVQYGEGERNGLYADNDNDSTFTGFLLYHD 229
```

RESULT 8

US-08-463-911-4

; Sequence 4, Application US/08463911

; Patent No. 5869330

; GENERAL INFORMATION:

; APPLICANT: Scherer, Philipp E.

; APPLICANT: Lodish, Harvey F.

; TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED

; TITLE OF INVENTION: EXCLUSIVELY IN ADIPOCYTES

; NUMBER OF SEQUENCES: 7

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.

; STREET: Two Militia Drive

; CITY: Lexington

; STATE: Massachusetts

; COUNTRY: USA

; ZIP: 02173

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/463,911

; FILING DATE:

; CLASSIFICATION: 530

; ATTORNEY/AGENT INFORMATION:

; NAME: Granahan, Patricia

; REGISTRATION NUMBER: 32,227

; REFERENCE/DOCKET NUMBER: WHI95-05

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (617) 861-6240

; TELEFAX: (617) 861-9540

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 246 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-08-463-911-4

Query Match

Best Local Similarity 24.1%; Score 316; DB 2; Length 246;

Matches 84; Conservative 38; Mismatches 85; Indels 56; Gaps 9;

```
QY 1 MWLLLVATPLLVHSSRGPAHYEMLGRWCDPHGPRGPDG-----APASVPPFP-- 53
Db 15 LLLLFLLALPLRSASAG--CYGIPGM-----PGMPGAPKGDHDLGQGPKEGPIAV 66
QY 54 PGAKEVRRGKAGLRGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 112
Db 67 PCTQGPKGKGPFGPHGRKNGPRGTSGLPDPGPRGPPGPPGVEG----- 113
```

```
QY 113 YAGLRPH-----EGYEVLRFDVVTVNNGNAYEAASGKTCMPGYYFFAY 158
Db 114 --RYQKHQGSVFTVTRQTTQYPEANALVRFNSVTVNPOGHYNPSTGKETCEVPLGYFYV 171
QY 159 HVLMRGGDGTSMWADLMKNGQVRASIAQADADQNYD---YASNSVILHLDVGEYFIKLD 215
Db 172 Y-----TSTANLCVHLNLTARVASFCDHMFNKSQVSSGGLALLRQGRDGVWLSYN 223
QY 216 GGVHGGNTNKYSTFSGFIYPD 238
Db 224 DINGWVGIEGNSVFSGFLFFD 246
```

RESULT 9

US-09-140-804-4

; Sequence 4, Application US/09140804

; Patent No. 6197930

; GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Humes, Jacqueline M.

; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS

; FILE REFERENCE: 97-49

; CURRENT APPLICATION NUMBER: US/09/140,804

; CURRENT FILING DATE: 1998-08-26

; EARLIER APPLICATION NUMBER: 60/056,983

; EARLIER FILING DATE: 1997-08-26

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 4

; LENGTH: 245

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-140-804-4

Query Match 24.0%; Score 314; DB 4; Length 245;

Best Local Similarity 36.6%; Pred. No. 1.3e-19;

Matches 87; Conservative 23; Mismatches 80; Indels 48; Gaps 11;

```
QY 28 CRMVCDPHGPRG-----PGPDGAPASVPPFPFGAKGEVG-----RRGKAGLRGPPGPPG 76
Db 26 CRA---PDGKKGEAGRPGRRGR-----PGLKGQGEPPGAPGIRTGIOGLKGDQGEPP 74
QY 77 PRGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 129
Db 75 PSGNPKGVGPGSGPLGARGIPGKGTGSPGNKQDPRPAFSAIRRNPPMGNNVIFD 134
QY 130 DVTVNNGNAYEAASGKTCMPGYYFFAYHVLMRGGDGTSMWADLM-----KNGQVRASA 184
Db 135 TVITNQEEPYQNHSGRFVCTVPGYYFTFQVL-----SQWEICLSIVSSSRGQVRRSL 187
QY 185 IAQDADQN--YDYASNSVILHLDVGEYFIKLDGKGVH---GGNTNKYSTFSGFIYYP 237
Db 188 GFCDITNKGIFQVWSSGMYLQLQQGDQVWVKDPKKGHIYQGSSEAD--SVFSGFLIFP 243
```

RESULT 10

US-09-140-804-8

; Sequence 8, Application US/09140804

; Patent No. 6197930

; GENERAL INFORMATION:

; APPLICANT: Sheppard, Paul O.

; APPLICANT: Humes, Jacqueline M.

; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS

; FILE REFERENCE: 97-49

; CURRENT APPLICATION NUMBER: US/09/140,804

; CURRENT FILING DATE: 1998-08-26

; EARLIER APPLICATION NUMBER: 60/056,983

; EARLIER FILING DATE: 1997-08-26

; NUMBER OF SEQ ID NOS: 47

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 8

; LENGTH: 247

; TYPE: PRT

; ORGANISM: Homo sapiens
US-09-140-804-8

Query Match 23.6%; Score 309.5; DB 4; Length 247;
Best Local Similarity 34.4%; Pred. No. 3.3e-19;
Matches 89; Conservative 32; Mismatches 95; Indels 43; Gaps 10;

QY 3 LLLVAIPLLVHSSRGAHYEMIGRCRMVCDPHG-----PRGPGDGPAPASVPPFP 53
Db 7 LFLFLILP--SHAEDDVTTTELAPA-LVPPPKGTCAGWMAGIPGPHGNGTTPGR--DGR 61

QY 54 PCAKEVGRRKAGLRGP-----PGPPGPRGPPGPPGPPGPPGPPGPPGPPGPPG 104
Db 62 DGTPEKEGKGDAGLLGPKGTGDMGTGAEGPRGFPQTGPKGEPG-----EA 110

QY 105 GYVPRIAFYAGLR-RPHEGYEVLRFDDVTVNNGNAYEASGKFTCPMPGVFFAYHVLMR 163
Db 111 AYMYRSASFVGLTRVTVNPNVPIRFTKIFYNQNHYDGTGKFCYCNIPGLYFYSHTVY 170

QY 164 GGDGTSMDWMLKMGQVRASAIADQADQNYDASNSVILHLDVGDVEFIKLDGKVGHG- 222
Db 171 MKD---VKVSLFKDKAVLFYTDYQYQKENVQASGVLLHLEVGQVWLQVYGDGHNGL 227

QY 223 ---NTNKYSTSGFIIPD 238
Db 228 YADNVND-STFTGFLLYHD 245

RESULT 11

US-09-118-408-3

; Sequence 3, Application US/09118408A
; Patent No. 6265544
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-30
; CURRENT APPLICATION NUMBER: US/09/118,408A
; EARLIER FILING DATE: 1998-07-17
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-118-408-3

Query Match 23.6%; Score 309.5; DB 4; Length 247;
Best Local Similarity 34.4%; Pred. No. 3.3e-19;
Matches 89; Conservative 32; Mismatches 95; Indels 43; Gaps 10;

QY 3 LLLVAIPLLVHSSRGAHYEMIGRCRMVCDPHG-----PRGPGDGPAPASVPPFP 53
Db 7 LFLFLILP--SHAEDDVTTTELAPA-LVPPPKGTCAGWMAGIPGPHGNGTTPGR--DGR 61

QY 54 PCAKEVGRRKAGLRGP-----PGPPGPRGPPGPPGPPGPPGPPGPPGPPGPPG 104
Db 62 DGTPEKEGKGDAGLLGPKGTGDMGTGAEGPRGFPQTGPKGEPG-----EA 110

QY 105 GYVPRIAFYAGLR-RPHEGYEVLRFDDVTVNNGNAYEASGKFTCPMPGVFFAYHVLMR 163
Db 111 AYMYRSASFVGLTRVTVNPNVPIRFTKIFYNQNHYDGTGKFCYCNIPGLYFYSHTVY 170

QY 164 GGDGTSMDWMLKMGQVRASAIADQADQNYDASNSVILHLDVGDVEFIKLDGKVGHG- 222
Db 171 MKD---VKVSLFKDKAVLFYTDYQYQKENVQASGVLLHLEVGQVWLQVYGDGHNGL 227

QY 223 ---NTNKYSTSGFIIPD 238
Db 228 YADNVND-STFTGFLLYHD 245

RESULT 12

US-09-506-855-3
; Sequence 3, Application US/09506855
; Patent No. 6448221
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; APPLICANT: Lasser, Gerald W.
; APPLICANT: Bishop, Paul D.
; TITLE OF INVENTION: INHIBITORS FOR USE IN HEMOSTASIS AND
; FILE REFERENCE: 99-12
; CURRENT APPLICATION NUMBER: US/09/506,855
; CURRENT FILING DATE: 2000-02-17
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-506-855-3

Query Match 23.6%; Score 309.5; DB 4; Length 247;
Best Local Similarity 34.4%; Pred. No. 3.3e-19;
Matches 89; Conservative 32; Mismatches 95; Indels 43; Gaps 10;

QY 3 LLLVAIPLLVHSSRGAHYEMIGRCRMVCDPHG-----PRGPGDGPAPASVPPFP 53
Db 7 LFLFLILP--SHAEDDVTTTELAPA-LVPPPKGTCAGWMAGIPGPHGNGTTPGR--DGR 61

QY 54 PCAKEVGRRKAGLRGP-----PGPPGPRGPPGPPGPPGPPGPPGPPGPPGPPG 104
Db 62 DGTPEKEGKGDAGLLGPKGTGDMGTGAEGPRGFPQTGPKGEPG-----EA 110

QY 105 GYVPRIAFYAGLR-RPHEGYEVLRFDDVTVNNGNAYEASGKFTCPMPGVFFAYHVLMR 163
Db 111 AYMYRSASFVGLTRVTVNPNVPIRFTKIFYNQNHYDGTGKFCYCNIPGLYFYSHTVY 170

QY 164 GGDGTSMDWMLKMGQVRASAIADQADQNYDASNSVILHLDVGDVEFIKLDGKVGHG- 222
Db 171 MKD---VKVSLFKDKAVLFYTDYQYQKENVQASGVLLHLEVGQVWLQVYGDGHNGL 227

QY 223 ---NTNKYSTSGFIIPD 238
Db 228 YADNVND-STFTGFLLYHD 245

RESULT 13

US-09-336-536-4
; Sequence 4, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-336-536-4

Query Match 22.9%; Score 300; DB 4; Length 228;
Best Local Similarity 35.9%; Pred. No. 1.9e-18;
Matches 79; Conservative 27; Mismatches 80; Indels 34; Gaps 7;

QY 40 PGDGPAPASVPPFPPCAKEVGRRKAGLRGPPGPPGPPGPPGPPGPPGPPGPPGPPG 95
Db 14 PCHPLPGT--PGHHGSQGLPGRDGRDGRDGRDGRDGRDGRDGRDGRDGRDGRDGRDGR 71

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2003, 16:46:45 ; Search time 21 Seconds
(without alignments)
1147.202 Million cell updates/sec

Title: US-10-005-499-378

Perfect score: 1311

Sequence: 1 MVLILLVAIPLLVHSSRGPA.....VHGNTKNYTFSGFIIPD 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 383519 seqs, 10123694 residues

Total number of hits satisfying chosen parameters: 383519

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA.*

- 1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB_PEP.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB_PEP.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB_PEP.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB_PEP.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB_PEP.*
- 6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB_PEP.*
- 7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB_PEP.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB_PEP.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB_PEP.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB_PEP.*
- 11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB_PEP.*
- 12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB_PEP.*
- 13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB_PEP.*
- 14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB_PEP.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	50.8	268	9	US-09-998-582-2
2	662.5	50.5	250	9	US-09-998-563-2
3	376	28.7	744	9	US-10-171-311-40
4	345.5	26.4	638	9	US-10-001-887-108
5	345.5	26.4	703	9	US-10-219-449-4
6	345.5	26.4	717	9	US-10-219-449-2
7	327	24.9	285	9	US-09-866-050A-382
8	327	24.9	294	9	US-09-866-050A-294
9	319.5	24.4	289	9	US-09-866-050A-630
10	318	24.3	244	9	US-09-911-176B-48
11	318	24.3	244	9	US-10-180-762-51
12	318	24.3	244	10	US-09-776-976-6
13	318	24.3	244	10	US-09-758-055-6
14	318	24.3	244	10	US-09-909-547-6
15	317.5	24.2	247	10	US-09-776-976-4
16	317.5	24.2	247	10	US-09-758-055-4
17	317.5	24.2	247	10	US-09-909-547-4
18	316	24.1	245	10	US-09-800-729-116
19	314	24.0	245	9	US-09-911-176B-49

20	314	24.0	245	9	US-10-180-762-52	Sequence 52, Appl
21	314	24.0	245	10	US-09-789-561-154	Sequence 154, Appl
22	314	24.0	245	10	US-09-800-729-85	Sequence 85, Appl
23	314	24.0	245	10	US-09-800-729-115	Sequence 115, Appl
24	314	24.0	245	10	US-09-800-729-118	Sequence 118, Appl
25	314	24.0	245	10	US-09-800-729-119	Sequence 119, Appl
26	311	23.7	229	10	US-09-800-729-117	Sequence 117, Appl
27	309.5	23.6	247	9	US-09-911-176B-3	Sequence 3, Appl
28	309.5	23.6	247	9	US-10-180-762-3	Sequence 3, Appl
29	309.5	23.6	247	9	US-10-241-258-3	Sequence 3, Appl
30	306	23.3	247	10	US-09-776-976-2	Sequence 2, Appl
31	306	23.3	247	10	US-09-758-055-2	Sequence 2, Appl
32	306	23.3	247	10	US-09-909-547-2	Sequence 2, Appl
33	300	22.9	243	9	US-09-944-413-42	Sequence 42, Appl
34	300	22.9	243	9	US-09-944-403-42	Sequence 42, Appl
35	300	22.9	243	9	US-09-944-896-42	Sequence 42, Appl
36	300	22.9	243	9	US-09-944-944-42	Sequence 42, Appl
37	300	22.9	243	9	US-09-944-907-42	Sequence 42, Appl
38	300	22.9	243	9	US-09-944-929-42	Sequence 42, Appl
39	300	22.9	243	9	US-10-028-072-362	Sequence 362, Appl
40	300	22.9	243	9	US-10-121-049-362	Sequence 362, Appl
41	300	22.9	243	9	US-10-123-904-362	Sequence 362, Appl
42	300	22.9	243	9	US-10-140-470-362	Sequence 362, Appl
43	300	22.9	243	9	US-09-796-753-68	Sequence 68, Appl
44	300	22.9	243	9	US-10-175-746-362	Sequence 362, Appl
45	300	22.9	243	9	US-10-176-918-362	Sequence 362, Appl

ALIGNMENTS

RESULT 1

US-09-998-582-2
; Sequence 2, Application US/09998582
; Patent No. US2002016047A1
; GENERAL INFORMATION:
; APPLICANT: Fox, Brian
; TITLE OF INVENTION: ADIPOCYTE COMPLEMENT RELATED PROTEIN
; TITLE OF INVENTION: ZACRP11
; FILE REFERENCE: 00-91
; CURRENT APPLICATION NUMBER: US/09/998,582
; CURRENT FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: US 60/253,863
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-998-582-2

Query Match	50.8%	Score	666;	DB	9;	Length	268;
Best Local Similarity	58.8%	Pred. No.	1.5e-38;				
Matches	144;	Conservative	15;	Mismatches	40;	Indels	46;
Gaps	10;						
QY	1	MVLLLLVAIPLLVHSSRGPAHYEMLGRCRMWCDPH-GPRGPG	---	PD-GAPASVPPPP	54		
Db	1	MVLLLLVILPVLSSAGTSAHYEMLGTCRMWCDPTGGTKAPSTAATPDRLGMLQSLPTFIQ	60				
QY	55	GAKGEVGRGKAGLRG	---	PGPPGPRGP	---	GEFGRPGPPGPGGVPAAAGYVP	108
Db	61	GPKEAGRGKAGLRG	---	PGPPGPGPPGPGGVPAAAGYVP	108		
QY	109	RIAFYAGLRPRHGEYEVLRFDVVTVNVCNAYEAA	---	SGKFTCPMPGVFFFAHYHLMRG	165		
Db	109	-----GLNAL-----VSRHVTNTVDACLFDSRI--PIPGIYFFTYQLMRGG	149				
QY	166	DGTSWADLMKNGQVRASIAODADQNDYASNSVILHLDVGDVEFIKLDGKVKHGGNTN	225				
Db	150	DGTSWADLCKNNQVRASIAODADQNDYASNSVILHLDVGDVEFIKLDGKVKHGGNTN	209				
QY	226	KYSTF	230				

[illegible]

Db 201 ITLA---NKHLAIGLVHNGQYRIPTF--DANTGNHVDVSGSTILALKEGDEVWLQIFVSE 255
QY 219 VHGGNTNKY---STFGFTIYPD 238
Db 256 QNGLFYDPYWTDSLFTGFLIYAD 278.

RESULT 8

US-09-866-050A-294
; Sequence 294, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 294
; LENGTH: 294
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-294

Query Match 24.9%; Score 327; DB 9; Length 294;
Best Local Similarity 32.7%; Pred. No. 2.4e-15;
Matches 86; Conservative 33; Mismatches 80; Indels 64; Gaps 9;

QY 26 GRCRWVCDPHGPRG-PGPDGAPASV-----PPPPGGA 56
Db 39 GGPQLVCSLPGPGPPGPGGSGMVRMGFPKDGODGDRGSGEGPGRTGN 98
QY 57 KGEVRRRGKAGLRGPPGPRGPPGPPGPPG-----GPGVAPAAAGVPRIA 111
Db 99 RCKQGPKGAGAGRAGPRGKGVSGTPCKHPIKGGKKGKGLPGPCS-----151
QY 112 FYAGLRPPHEGEV-----LRPDDVVTVNGNAYEAASGKFTCPMPGVFFAYH 159
Db 152 --CGSSRAKSAFSAVTRKSYPRERLPKPKILMNEGGHYNASSGKFVCSVPGIYFTYD 209
QY 160 VLMRGDGTSMWADLMKNQVRSALAQDADQ--NYDYASNSVILHLDVGDVEFIKLDGK 218
Db 210 ITLA---NKHLAIGLVHNGQYRIPTF--DANTGNHVDVSGSTILALKEGDEVWLQIFVSE 264
QY 219 VHGGNTNKY---STFGFTIYPD 238
Db 265 QNGLFYDPYWTDSLFTGFLIYAD 287

RESULT 9

US-09-866-050A-630
; Sequence 630, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 630
; LENGTH: 289
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-630

Query Match 24.4%; Score 319.5; DB 9; Length 289;
Best Local Similarity 32.6%; Pred. No. 7.6e-15;
Matches 92; Conservative 35; Mismatches 104; Indels 51; Gaps 12;
QY 2 VILLVVAIPLLVHSSRGPAHYEMLGRC-----RMVCD--PHGPRGPGPDGAPASVPPF-----52
Db 1 MIVLLYVTSLAICASQGRGNQAKGSYSPRVICISPLGLPGPPGPGANGSPGPHGRIGL 60
QY 53 -----PPGAKGEVGRGKAGLRGPPG-----PPRGPGEGRPPGPPG 92
Db 61 PGDRGDRGRKGEKGTAGLKGKGLAGLAGEKQDGETGKKGPIGPEGEKEGEVGPAG 120
QY 93 PGP-----GPG--GVAPAAAGVPRIAFYAGLR--PHEGYEVLRFDDVTVNGNAYE 140
Db 121 PGPKGDRGDQDGLPGVCRGSIVLKSAFSVGITTSYPERLPIL--FNKVLNNEGEHYN 179
QY 141 AASGKFTCPMPGVFFAYHVMRGDGTSMWADLMKNQVRSALAQDADQ--NYDYASNS 199
Db 180 PATGKFICAFPGIYFYSYDITLA---NKHLAIGLVHNGQYRIPTF--DANTGNHVDVSGS 234
QY 200 VILHLDVGDVEFIKL---DGGKVHGGNTNKYSTFGFTIYPD 238
Db 235 TVIYLQPEDEVWLEIFFNDQNGLFSDPGWADSLFSGFLIYD 276

RESULT 10

US-09-911-176B-48
; Sequence 48, Application US/09911176B
; Patent No. US20020156243A1
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ANTIBODIES THAT BIND AN
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOG
; FILE REFERENCE: 97-30D1
; CURRENT APPLICATION NUMBER: US/09/911,176B
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: 09/118,408
; PRIOR FILING DATE: 1998-07-17
; PRIOR APPLICATION NUMBER: 60/053,154
; PRIOR FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 48
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-911-176B-48

Query Match 24.3%; Score 318; DB 9; Length 244;
Best Local Similarity 34.3%; Pred. No. 8.2e-15;
Matches 92; Conservative 36; Mismatches 82; Indels 58; Gaps 12;
QY 3 LLLVVAIPLLVHSSRGPAHYEMLGRRCRWVCDPHGPRGPG-----PDGA-----PASVPPPP 53
Db 1 MLLLGAVLLLLAL---PGHDQ-----ETTTQGGVLLPLPKGACTGWMAGIPGHP 47
QY 54 -----PGAKGEVRRGKAGLRGPPGPPGPRGPPGPRGPPGPPG-----PGPG 98
Db 48 GHNGAPGRDGRDGTGEGKEKGDPLGKIDIGETGVCAGGPRGPPGPGIQRKGGEPGG 107
QY 99 GVAPAAGVPRIAFYAGLRPPHEGEV-----LRPDDVVTVNGNAYEAASGKFTCPMPGV 153
Db 108 -----AYVRSASFVGL---ETVVTIPNMPRETKIFYNQNNHYDGTGKFKHCNIPGL 157
QY 154 YFFAYHVMRGDGTSMWADLMKNQVRSALAQDADQ--NYDYASNSVILHLDVGDVEFIK 213

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2003, 16:43:09 ; Search time 18 Seconds
(without alignments)
1271.111 Million cell updates/sec

Title: US-10-005-499-378

Perfect score: 1311

Sequence: 1 MVLLLVLAIPLLVHSSRPA.....VHGNTNKYFTSGFIIPD 238

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_73.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	389	29.7	743	1	collagen alpha 1(V
2	388.5	29.6	744	1	collagen alpha 1(V
3	383.5	29.3	744	1	collagen alpha 1(V
4	376	28.7	680	2	collagen alpha 1(X
5	376	28.7	744	2	collagen alpha 1(X
6	369.5	28.2	674	2	collagen alpha 1(X
7	362.5	27.7	680	1	collagen alpha 1(X
8	360	27.5	674	2	collagen alpha 1(X
9	352	26.8	253	1	complement subcomp
10	345	26.3	635	2	collagen alpha 2(V
11	322	24.6	245	1	complement subcomp
12	318	24.3	244	2	gelatin-binding 28
13	316	24.1	246	2	complement subcomp
14	316	24.1	253	2	complement subcomp
15	314	24.0	245	1	complement subcomp
16	305.5	23.3	253	2	complement C1q B c
17	302	23.0	245	2	complement subcomp
18	300	22.9	219	2	hypothetical prote
19	268.5	20.5	215	2	hibernation-relate
20	261.5	19.9	215	2	hibernation-relate
21	260.5	19.9	992	2	hypothetical prote
22	250	19.1	423	2	collagen precursor
23	243	18.5	224	2	cerebellin-like gl
24	229	17.5	193	2	cerebellin precurs
25	222.5	17.0	196	2	hibernation-relate
26	221.5	16.9	464	2	collagen II Al pro
27	219.5	16.7	1486	1	collagen alpha 1(I
28	214.5	16.4	473	2	collagen - Chicken
29	211.5	16.1	1464	1	collagen alpha 1(I

30	211	16.1	1466	1	CGHU7L	collagen alpha 1(I
31	210.5	16.1	1492	1	A40333	collagen alpha 1'(
32	208.5	15.9	1042	1	CGCH1S	collagen alpha 1(I
33	208	15.9	453	2	S18804	collagen alpha 4(I
34	207.5	15.8	1315	2	A56101	collagen alpha 1(X
35	207.5	15.8	1532	2	B56101	collagen alpha 1(X
36	207	15.8	1774	2	B56101	collagen alpha 1(X
37	206.5	15.8	1419	2	A41182	collagen alpha 1(I
38	206.5	15.8	1487	1	B41182	collagen alpha 1(I
39	205.5	15.7	779	1	CGBO1S	collagen alpha 1(I
40	205	15.6	677	2	S23296	collagen alpha 2(I
41	204	15.6	193	2	S07133	collagen alpha 1(I
42	204	15.6	394	2	T33641	hypothetical prote
43	203.5	15.5	170	2	B57131	collagen alpha 2(V
44	203.5	15.5	2944	2	A54849	collagen alpha 1(V
45	203	15.5	102	2	B34770	ORE2 protein - sai

ALIGNMENTS

RESULT 1

S23779

collagen alpha 1(VIII) chain - mouse

C:Species: Mus musculus (house mouse)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C:Accession: S23779

R.Muragaki, Y.; Shiota, C.; Inoue, M.; Ooshima, A.; Olsen, B.R.; Ninomiya, Y.

Eur. J. Biochem. 207, 895-902, 1992

A:Title: Alpha-1(VIII)-collagen gene transcripts encode a short-chain collagen polype

A:Reference number: S23779; MUID:92362626; PMID:1499564

A:Accession: S23779

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-743 <MUR>

A:Cross-references: EMBL:X66976; NID:g50493; PIDN:CAA47387.1; PID:g1359953

C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homolog

F:616-742/Domain: complement C1q carboxyl-terminal homology <CIQ>

Query Match 29.7%; Score 389; DB 1; Length 743;

Best Local Similarity 38.1%; Pred. No. 5e+21; Mismatches 30; Indels 46; Gaps 10;

Matches 94; Conservative 77; Indels 46; Gaps 10;

QY	31	VCDPHGPRG----	PGDGPAPASVPPPP	PGAKGEVGRGKAGLRG	PPGPPGPPGPPGPPGPR 86
Db	502	IVGPSGPIGPGIPGK	GEPL--PGPGFPG-VGK	PGVAGLHGPGEKALGPQ	558
QY	87	PGPPGPPG-PGPGV---	APAAAGVVPRIAF-----	YAGLRPHEG--YEV----	125
Db	559	PGPPGPPGPPGPPAY	MTPTSPQGEYLPDMGLG	IDGVKTPHAYAGKKHGG	PAYEMPAFT 618
QY	126	-----	LRDDVVTVNGNAYEAS	CKTCMPGVYFFAYHVL	MRGGDGTSMWA 172
Db	619	AELTVPPFPPGAPVK	FDKLLYNGRQNYNPQT	GIFCEVPGVYFFAYHVK	CKGG---NVWV 675
QY	173	DMKNGQVRSAAIAQ	ADQNDYASNVILHLVD	GEVFIKIDGKGVHGGNT	KY--STF 230
Db	676	ALFKNNEPMYTYDEY	KKGFLDQASGSAVLL	LRPGDQVFLQNPFEQA	RLYAGQYVHSSF 735
QY	231	SGFIYP 237			
Db	736	SGYLLYP 742			

RESULT 2

A34246

collagen alpha 1(VIII) chain precursor - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: A34246

R.Yamaguchi, N.; Benya, P.D.; van der Rest, M.; Ninomiya, Y.

J. Biol. Chem. 264, 16022-16029, 1989

A:Title: The cloning and sequencing of alpha1(VIII) collagen cDNAs demonstrate that t


```

QY   109 RTAFYA-----GLRRPHEGYEVLREDVVTVNGNAYEAASGKFTCPMPGVYFAFHVLM 162
      :||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| |
Db    121 KIAFSATRTINPLRRD----QTIREFDHVTNNNNYEPRSGKFTCKVPGLYYFTVHASS 176
      :||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| |
QY   163 RGGDGTSMADLMKNQVRASATAQAADQND---YASNVSILHLVDGDSEVFKLDDGKV 219
      :||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| |
Db    177 RG-----NLCVNLMR-GRERAQKVVTFCDYAYNTFQVTTTGGVMVLKLEGENVENFLQATDKNS 231
      :||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| |
QY   220 HGGNTNKYSTFGSFIIPD 238
      :||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| |
Db    232 LLIGMEGANISFSGFLIFPD 250
      :||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| |

RESULT 10
A57131
collagen alpha 2(VIII) chain - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 03-Nov-1995 #sequence_revision 03-Nov-1995 #text_change 13-Aug-1999
C:Accession: A57131
R:Muragaki, Y.; Jacenko, O.; Apté, S.; Mattei, M.G.; Ninomiya, Y.; Olsen, B.R.
J. Biol. Chem. 266, 7721-7727, 1991
A:Title: The alpha2(VIII) collagen gene. A novel member of the short chain coll
A:Reference number: A57131; MUID:91210292; PMID:2019595
A:Accession: A57131
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-635 <MUR>
A:Cross-references: GB:M60832; NID:g177178; PIDN:AAA62822.1; PID:g177179
C:Genetics:
A:Gene: GDB:COL9A2
A:Cross-references: GDB:127812; OMIM:120252
A:Map position: lp34.3-lp32.3
C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal h
F:1-11/Domain: amino-terminal nonhelical (fragment) #status predicted <NC2>
F:12-468/Region: interrupted helical
F:469-635/Domain: carboxyl-terminal nonhelical #status predicted <NC7>
F:508-634/Domain: complement C1q carboxyl-terminal homology <CIQ>

Query Match          26.3%; Score 345; DB 2; Length 635;
Best Local Similarity 33.0%; Pred. No. 7e-18;
Matches 96; Conservative 32; Mismatches 67; Indels 96; Gaps 13;

QY   35 HGPRGP-GPDGAPA-SVPPFPPGA-----KGVEGRGKAGLRGP----- 71
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    352 HGGPGPTGPKGEFTGRPGCGPVAGALGQKDGLGPLQGPGLRGSPGIGLQGPAGPIGP 411
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY   72 -----PGPP-----GPRGPGEGRCPGPPGPG--PGPGGVAPAAQYV 107
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    412 QGLPLGLKEGLPGPEGRAGEPGTAGRPGEVPGVPSGITGPGPLGPPG-APGAFFE 470
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY   108 PRTAFYAGRLRRPHEGYE-----GELSKKKKPGQFGLGELSAAHTPAFTAVLTSPASGMPVKFD 129
      :||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| |
Db    471 TGI---AGLHLPGVGEGAVLGKGGKPGQFGLGELSAAHTPAFTAVLTSPASGMPVKFD 527
      :||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| |
QY   130 DVVTVNGNAYEAASGKFTCPMPGVYFAFHYVLMRGDGTSMWADLMKNQVRASAIQA 189
      :||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| |
Db    528 RTLYNGHSYNPATGIFTCPGVGYFYAHVHK---GTNVVALYKN-NVPATYTYDEV 583
      :||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| |
QY   190 DQNY-DYASNVSILHLVDGDSEVFKLDDGKVHGNTNKY--STFGSFIIP 237
      :||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| |
Db    584 KKGYLDQASGGAVALQLRPNDQVWPVPSDAQNGLYSTEYIHSSFGSLCLCP 634
      :||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| | : ||| |

RESULT 11
C1HUQC
complement subcomponent C1q chain C precursor - human
N:Alternate names: complement subcomponent C1q gamma chain
C:Species: Homo sapiens (man)
C>Date: 22-May-1981 #sequence_revision 31-May-1996 #text_change 22-May-1998
C:Accession: S14351; A03207
R:Sellar, G.C.; Blake, D.J.; Reid, K.B.M.
Biochem. J. 274, 481-490, 1991

```

A; Experimental source: adipose tissue
R; Nakano, Y.; Tobe, T.; Choi-Miura, N.H.; Mazda, T.; Tomita, M.
J. Biochem. 120, 803-812, 1996

A; Title: Isolation and characterization of GBP28, a novel gelatin-binding protein pnpA
A; Reference number: JC4944; MUID: 97103474; PMID: 8947845

A; Accession: JC4944

A; Molecule type: protein

A; Residues: 19-38; 93-100; 101-112; 135-149; 173-178 <NAK>

C; Comment: This protein is an endogenous factor that binds with a collagen-like domain.

C; Genetics:

A; Gene: apm1

C; Superfamily: unassigned collagens; complement Clq carboxyl-terminal homology

C; Keywords: adipose tissue; glycoprotein; hydroxyproline

F; 1-18/Domain: signal sequence #status predicted <Sig>

F; 19-244/Product: gelatin-binding 28kDa protein #status experimental <MAT>

F; 42-107/Region: collagen-like

F; 114-241/Domain: complement Clq carboxyl-terminal homology <ClQ>

F; 95/Modified site: 4-hydroxyproline (Pro) #status experimental

F; 230/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.3%; Score 318; DB 2; Length 244;
Best Local Similarity 34.3%; Pred. No. 2.4e-16;
Matches 92; Conservative 36; Mismatches 82; Indels 58; Gaps 12;

QY 3 LLLLVAILPLVHSSRGPAHYEMLGRVMWCDPHGPRGP-----PDGA----PASVPPFP 53
:
DB 1 MLLLGAVLLLLAL--PGHDQ-----ETTQGGVLLPLPKGACTGMAGIPCHP 47

QY 54 -----PGAKEVGRRGAKGLRGPPGPGPGPEGRGPGPPG-----PGPG 98
:
DB 48 GHNGAPGRDRDGTPEGKGKDGPLGIKPGIDIGETVPGEAGRPFGIQRKGEPGE 107

QY 99 GVAPAAGYVPRIAFYACLRRLRPHGEYEV-----LRFDDVTNVGNAYAAASCKFCMPGV 153
:
DB 108 -----AYVRSAFSGVL---ETYVTIPNPIRFTKIFYNQNHYGSTGKFHCNIPGL 157

QY 154 YFFAYHLMRGDGTSMADLMKNQVRSAAIAODADONDYASNVSILHLVDGDEVFIK 213
:
DB 158 YYFAYHTVTMKD---VKVSLFKDKKAMLFYDQYQENNVDAQSGSVLLHLEVGDOVWLQ 214

QY 214 LDG-GKVHG--GNTNKYSTSGFTIYPD 238
:
DB 215 VYGEGERNGLYANDNDSTFTGFLLYHD 242

RESULT 13
S29328
Complement subcomponent Clq chain C - mouse
C; Species: Mus musculus (house mouse)
C; Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 20-Aug-1999
C; Accession: S29328
R; Petry, F.; Reid, K.B.M.; Loos, M.
Eur. J. Biochem. 209, 129-134, 1992

A; Title: Isolation, sequence analysis and characterization of cDNA clones coding for ecerbellin.

A; Reference number: S29328; MUID: 93011118; PMID: 1396691

A; Accession: S29328

A; Status: preliminary

A; Molecule type: mRNA

A; Residues: 1-246 <PET>

A; Cross-references: EMBL: X66295; NID: g50228; PIDN: CA446993.1; PID: g50229

C; Superfamily: complement subcomponent Clq chain A; complement Clq carboxyl-terminal

F; 122-245/Domain: complement Clq carboxyl-terminal homology <ClQ>

Query Match 24.1%; Score 316; DB 2; Length 246;
Best Local Similarity 31.9%; Pred. No. 3.4e-16;
Matches 84; Conservative 38; Mismatches 85; Indels 56; Gaps 9;

QY 1 MWILLVAIPLLVHSSRGPAHYEMLGRVMWCDPHGPRGPCPDG-----APASVPPFP-- 53
:
DB 15 LLLLFLALPLSOASAG--CYGIPGM-----PGMGPAGKGDHDLQGKGEGPIPAV 66

QY 54 PGAKEVGRRGKAGLRCPGPPGPRGPPGPRGPPGPGPG--PGPGGVAPAAGYVPIATF 112

THIS PAGE BLANK (USPTO)

Result No.	Query %			DB	ID	Description
	Score	Match	Length			
1	584	44.5	120	11	Q8R1Z2	Q8rlz2 mus musculus
2	397	30.3	744	11	Q9D2V4	Q9d2v4 mus musculus
3	397	30.3	744	11	Q921S8	Q921s8 mus musculus
4	376	28.7	744	4	Q96D07	Q96d07 homo sapien
5	363	27.7	675	6	Q9N178	Q9n178 sus scrofa
6	362.5	27.7	295	11	Q921K4	Q921k4 rattus norv
7	345.5	26.4	705	4	Q8TEJ5	Q8tej5 homo sapien
8	334.5	25.5	240	6	Q95MQ4	Q95mq4 bos taurus
9	329	25.1	294	11	Q9P8U4	Q9p8u4 mus musculus
10	316	24.1	243	6	Q95JD7	Q95jd7 macaca mula
11	302	23.0	245	11	Q9DCM6	Q9dcm6 mus musculus
12	295	22.5	243	11	Q8R002	Q8r002 mus musculus
13	292	22.3	246	11	Q9ES30	Q9es30 mus musculus
14	282	21.5	194	6	Q95J95	Q95j95 canis famil
15	260.5	19.9	347	4	Q961H6	Q961h6 homo sapien
16	260.5	19.9	583	4	Q96G58	Q96g58 homo sapien


```
SQ SEQUENCE 744 AA; 73364 MW; 2BC1B0955DE2C9A3 CRC64;
Query Match
Best Local Similarity 37.2%; Score 376; DB 4; Length 744;
Matches 90; Conservative 31; Mismatches 77; Indels 44; Gaps 10;

Qy 34 PHGPRG--PGDGPAPASVPPPPGAKGVGRGKAGLRGPPGPPGPPGPPGPPGPP 92
D 11111111111111111111111111111111111111111111111111111
Db 508 PIGPPGIPGPKGEPGL--PGPPGPPG--IGKPGVAGLHGPCKGALGPGQGLGPPGP 564
Qy 93 PG-PGPGGV-----APAGVVPRIAF-YAGLRPH-----EGYEV----- 125
D 11111111111111111111111111111111111111111111111111111
Db 565 PGPPGPAPVPPPTPPGEXLPLDMGLGIDGVKPPHAYGAKKNGKNGPAYEMPATBELTA 624
Qy 126 -----LRFDDVVTVNGVNAEASGKFTCPMPGVFFAYHVLMRGGDGTSMWADLMKN 177
D 11111111111111111111111111111111111111111111111111111
Db 625 PFPVPGAPVFNKLLYNGRQNYNPQTGFTCEVGVFFAYHVKCKG---NWWVALFKN 681
Qy 178 GOVRASIAQADQNYDYASNSVILHLDVGDVEFIKLDGKGVHGGNTNKY--STFSGFII 235
D 11111111111111111111111111111111111111111111111111111
Db 682 NEPVMTYDEYKGFGLDQASGSAVLLLRPGDRVFLQMPSEQAAGLYAGQYVHSSFSGYLL 741
Qy 236 YP 237
D 11111111111111111111111111111111111111111111111111111
Db 742 YP 743

RESULT 5
Q9N178
ID Q9N178 PRELIMINARY; PRT; 675 AA.
AC Q9N178
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Type X collagen.
GN COL10A1.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN 11
SEQUENCE FROM N.A.
MEDLINE=21015405; PubMed=11130976;
RA Nielsen V.H., Bendixen C., Arnbjerg J., Sorensen C.M., Jensen H.E.,
RA Shukri N.M., Thomsen B.;
RT "Abnormal growth plate function in pigs carrying a dominant mutation
RT in type X collagen.";
RL Mamm. Genome 11:1087-1092(2000).
DR EMBL; AF222861; AAF37271.1; -.
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 8.
DR PRINTS; PR00007; COMPLEMNTC1Q.
DR PRODOM; PD000007; Collagen; 2.
DR SMART; SM00110; Clq; 1.
DR PROSITE; PS01113; Clq; 1.
KW Collagen.
SQ SEQUENCE 675 AA; 65447 MW; 26397B10310383F9 CRC64;

Query Match
Best Local Similarity 36.4%; Score 363; DB 6; Length 675;
Matches 87; Conservative 32; Mismatches 78; Indels 42; Gaps 9;

Qy 36 GPRG----PGPDG--APASVPPPPGAKGEVGR-----RGKAGLRGPPGPPGPR 78
D 11111111111111111111111111111111111111111111111111111
Db 441 GPRGAPGIPGTRGIPGPPGIPG--PGSKGDPGNGPPGAGIATKGLNGTPGPPGPK 499
Qy 79 GPPGEPGRPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 121
D 11111111111111111111111111111111111111111111111111111
Db 500 GHAGEPLGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 558
Qy 122 GYEV-LRFDDVVTVNGVNAEASGKFTCPMPGVFFAYHVLMRGGDGTSMWADLMKNQV 180
```

```
Db 559 AIGAPIPFDKILYNGQQHYDPKGTGFTCRIPGIYFYSYHIVK---GTHAVVGLYKNGTP 615
Qy 181 RASIAQADQNYDYASNSVILHLDVGDVEFIKLDGKGVHGGNTNKY--STFSGFIIYP 237
D 11111111111111111111111111111111111111111111111111111
Db 616 VMTYDEYVKGYLDOASGSAILDLDNDQVQLPNAGSNGLYSSEYVHSSFSGLVAP 674

RESULT 6
Q9Z1K4
ID Q9Z1K4 PRELIMINARY; PRT; 295 AA.
AC Q9Z1K4
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Collagen alpha 1 type X (Fragment).
GN COL10A1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN 11
SEQUENCE FROM N.A.
MEDLINE=20310874; PubMed=10853827;
RA Marks S.A., Lundmark C., Christersson C., Wurtz T., Odgren P.R.,
RA Selfert M.F., Mackay C.A., Mason-Savas A., Popoff S.E.;
RT "Endochondral bone formation in toothless (osteopetrotic) rats;
RT failures of chondrocyte patterning and type X collagen expression.";
RL Int. J. Dev. Biol. 44:309-316(2000).
DR EMBL; AJ31848; CAA10518.1; -.
DR InterPro; IPR001073; Clq.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 3.
DR PRINTS; PR00007; COMPLEMNTC1Q.
DR SMART; SM00110; Clq; 1.
DR PROSITE; PS01113; Clq; 1.
FT NON_TER 1
SQ SEQUENCE 295 AA; 30012 MW; FF43B1548028813E CRC64;

Query Match
Best Local Similarity 35.4%; Score 362.5; DB 11; Length 295;
Matches 87; Conservative 35; Mismatches 77; Indels 47; Gaps 10;

Qy 36 GPRG----PGPDG--APASVPPFP-----PGAKGEVG--RRKAGLRGPPGPPGPRG 79
D 11111111111111111111111111111111111111111111111111111
Db 52 GPRGEPGIPGTRGIPGPPGIPGPPGSKGDPKPGAPGAGIVTKGLNGPAGPPGPPGPRG 111
Qy 80 PGPGRGPPGPPGPPG--PGP-----GGVAPAAAGVVPRIAFYAGLR-RPHEGYEV 125
D 11111111111111111111111111111111111111111111111111111
Db 112 HTGEPGLGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 171
Qy 126 -----LRFDDVVTVNGVNAEASGKFTCPMPGVFFAYHVLMRGGDGTSMWAD 173
D 11111111111111111111111111111111111111111111111111111
Db 172 ILSKAYPAVGAPIPFDEILYNRQOHYDPRSGIFTCKIPGIYFYSYHIVK---GTHVWVG 228
Qy 174 LMKNGQVRASIAQADQNYDYASNSVILHLDVGDVEFIKLDGKGVHGGNTNKY--STFS 231
D 11111111111111111111111111111111111111111111111111111
Db 229 LYKNGTPTMTYDEYSKGYLDQASGSAIMELTENDQVQLPNAGSNGLYSSEYVHSSFS 288
Qy 232 GFIIYP 237
D 11111111111111111111111111111111111111111111111111111
Db 289 GFLVAP 294

RESULT 7
Q8TEJ5
ID Q8TEJ5 PRELIMINARY; PRT; 705 AA.
AC Q8TEJ5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
```

```
DE FLJ00201 protein (Fragment).
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=SPLEEN;
RA Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
RT spleen.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK074129; BAB84955.1; -
FT NON_TER
SQ SEQUENCE 705 AA; 67430 MW; 07DB85A65A948ED3 CRC64;

Query Match
Best Local Similarity 26.4%; Score 345.5; DB 4; Length 705;
Matches 94; Conservative 30; Mismatches 77; Indels 63; Gaps 12;

QY 17 RGAHYEMLGRMCRVCDPHGRGP-GPDCAPA-SVPPPPPPGAKGEVGRGKAGLRGPPGP 74
DB 461 RGPSPGIPGL-----QGAPGPGQGLKGLKGPGLGPPGE-GRAGEGTAGTGP 510
QY 75 PGPRGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 124
DB 511 PGVPSGPGTGPVPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 564
QY 125 -----VLREDDVVTVNGVAYEASGKFTCPMGVYFF 156
DB 565 QFGLGELSAHATPAFTAVLTSPFASGMPVKFDRLYNGHSGYNPAGTFTCPVGVVYIF 624
QY 157 AYHVLMRGGDGTSMWADLMKNGVRASAIADADQNY-DYASNSVILHLDVGDEVFIKLD 215
DB 625 AYHVHVK---GTVNWALYKN-NVPATYTYDEYKGYLDQASGGAVLQRPNDQVQVMP 680
QY 216 GGVHGGNTKNY--STFGSFIY 237
DB 681 SDQANGLYSTEYTHSFSGLLCP 704

RESULT 8
Q95M04 PRELIMINARY; PRT; 240 AA.
ID AC Q95M04;
DC 01-DEC-2001 (TrEMBLrel. 19, Created).
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Adipose tissue-specific protein adipon Q.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21369933; PubMed=11382781;
RA Sato C., Yasukawa Z., Honda N., Matsuda T., Kitajima K.;
RT "Identification and Adipocyte Differentiation-dependent Expression of
RT the unique Distal Acid Residue in an Adipose Tissue-specific
RT Glycoprotein, Adipo Q.";
RL J. Biol. Chem. 276:28849-28856(2001).
DR EMBL; AF269230; AAK58902.1; -
DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 1.
DR Prodom; PD000007; Collagen; 1.
DR PROSITE; PS01113; Clq; UNKNOWN.1.
SQ SEQUENCE 240 AA; 26091 MW; C6253BA803B9A668 CRC64;
```

Query Match 25.5%; Score 334.5; DB 6; Length 240;

```
Best Local Similarity 35.4%; Pred. No. 2.7e-20;
Matches 92; Conservative 32; Mismatches 79; Indels 57; Gaps 10;

QY 3 LLLLVAILLVHSSRGPAHVLMGRMCRVCDPHGRGP-----PGPDGAPASVP 50
DB 7 LLLLLALP-----SHGEDNNE-----DPLPKGACAGWAGIPLGPHGNGTGR-- 50
QY 51 PFPPGAKGEVRRGKAGLRG-----PGPPGPRGPPGPPGPPGPPGPPGPPGPPGPPG 101
DB 51 DGRDGTGPGKEGKGDAGLLGPKGETGVGMTGAEGPRGPGTTPGRKGEPG----- 100
QY 102 PAAGVYPRIAFYAGLR-RPHEGYEVLRFDDVTVNGVAYEASGKFTCPMGVYFFAYHV 160
DB 101 -EAAVYRSAFVSGLETRVTPNVPIRFTKIFYNQNHVDGSGTKFYCNIPGLYPSYHI 159
QY 161 LMRGGDGTSMWADLMKNGVRASAIADADQNYDVASNSVILHLDVGDEVFIKLDGKGVH 220
DB 160 TVYMKD---VKVSLFKKRAVLFTDYQYQEKVNDQASGVLLHLEVDQVWLQVYEGNH 216
QY 221 GG-----NTNKYSTFGSFIY 236
DB 217 NGVYADNVND-STFTGFLLY 235

RESULT 9
Q9D8U4 PRELIMINARY; PRT; 294 AA.
ID AC Q9D8U4;
DC 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE 1810033K05RIK protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=PANCREAS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK007683; BAB25187.1; -
DR MGD; MGI:1916433; 1810033K05RIK.
DR InterPro; IPR001073; Clq.
DR InterPro; IPR000087; Collagen.
DR Pfam; PF00386; Clq; 1.
DR Pfam; PF01391; Collagen; 2.
DR PRINTS; PR000007; COMPLEMENTC1Q.
DR SMART; SM00110; Clq; 1.
DR PROSITE; PS01113; Clq; 1.
SQ SEQUENCE 294 AA; 30865 MW; 6D3905AE7C19E6FA CRC64;
```

Query Match 25.1%; Score 329; DB 11; Length 294;

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)